ESSAYS IN MACRO-FINANCE AND ASSET PRICING

Amr Elsaify

A DISSERTATION

in

Finance

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

Nikolai Roussanov, Moise Y. Safra Associate Professor of Finance

Graduate Group Chairperson

Catherine Schrand, Celia Z. Moh Professor, Professor of Accounting

Dissertation Committee

Joao F. Gomes, Howard Butcher III Professor of Finance

Urban J. Jermann, Safra Professor of International Finance and Capital Markets

Nikolai Roussanov, Moise Y. Safra Associate Professor of Finance

ESSAYS IN MACRO-FINANCE AND ASSET PRICING

Dedicated to those without whom this would have been impossible.

You mean the world to me.

ACKNOWLEDGEMENT

I cannot put into words my gratitude towards my dissertation committee. The time, effort, and invaluable advice that Nikolai Roussanov, Joao Gomes, and Urban Jermann have given me is invaluable. I would also like to thank Andrew Abel, Anna Cororaton, Aycan Corum, Max Croce, Tetiana Davydiuk, Winston Dou, Vito Gala, Deeksha Gupta, Jessica Jeffers, Daniel Kim, Roni Kisin, Karen Lewis, Sophie Moinas, Christian Opp, Ryan Peters, Hongxun Ruan, Lin Shen, Luke Taylor, Jules van Binsbergen, Jessica Wachter, Ram Yamarthy, and Amir Yaron for helpful discussions and feedback. Grants from the Rodney White Center for Financial Research and the Jacobs Levy Institute Equity Mangement Center for Quantitative Financial Research are gratefully acknowledged.

ABSTRACT

ESSAYS IN MACRO-FINANCE AND ASSET PRICING

Amr Elsaify

Nikolai Roussanov

This dissertation consists of three parts. The first documents that more innovative firms earn higher risk-adjusted equity returns and proposes a model to explain this. Chapter two answers the question of why firms would choose to issue callable bonds with options that are always "out of the money" by proposing a refinancing-risk explanation. Lastly, chapter three uses the firm-level evidence on investment cyclicality to help resolve the aggregate puzzle of whether R&D should be procyclical or countercylical.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

PREFACE

This dissertation consists of three chapters. In the first, I show that firms that engage in innovative product development, as measured by the fraction of their investment that goes to Research and Development (R&D) activities, earn higher risk-adjusted equity returns. A portfolio that goes long the most innovative and shorts the least innovative firms earns a risk-adjusted return in excess of 7% per annum. R&D-intensive firms also tend to charge higher price markups. Combining insights from industrial organization with a production-based asset pricing framework, I propose a model in which heterogeneous firms produce vertically differentiated goods and market them to heterogeneous consumers. Firms are subject to aggregate demand and supply shocks, which are both priced by investors, and thus the return premium of innovative firms is explained by their differential exposures to these shocks. In addition to explaining this return spread, the model makes predictions on firm investments, future profit markups, and firm size that are consistent with the data.

The next paper (joint with Nikolai Roussanov) examines the surprising recent trend of U.S. corporations significantly increasing their issuance of callable debt in the past 10-15 years. Whereas callable debt was issued in the past for interest rate hedging motives, the vast majority of callable bonds issued today have call options that will never be "in the money". This feature implies that previous explanations for the issuance of callable debt no longer rationalize the current pattern. We present evidence on the types of firms issuing these bonds and their usage of the proceeds, which motivates a new theory for why firms desire these eternally "out of the money" call options. This theory captures the motives of firms in matching the maturities of investment and financing and endogenously generates firm-specific refinancing risk. We then embed this theory into a production-based model and show that callable bonds can expand access to capital markets and increase investment.

I study the joint effects of investment composition and leverage in the final chapter. Classical R&D theory implies that R&D expenditures should be countercyclical, but empirical

evidence finds that it is procyclical on aggregate. I reconcile these findings by documenting novel empirical evidence on how the composition of investment changes with firm profitability conditions. I show that R&D tends to lead to future increases in idiosyncratic volatility while Capex tends to lead to future increases in systematic volatility. For firms that are leverage constrained, this can limit their ability to finance R&D during bad times, leading to an observed procyclicality. I capture this effect in a simple model and show that it is consistent with the empirical evidence.

CHAPTER 1 : **The Innovation Premium**

1.1. **Introduction**

Research and development (R&D) activities are widely cited as a key driver of both firm-specific technological progress and aggregate economic growth. R&D is becoming an increasingly important investment activity for private corporations in particular, as firms shift from dependence on physical capital to intangible and knowledge capital (Congressional Budget Office (2005)). This shift is perhaps best illustrated by the change in investment composition: Figure 1 plots the cumulative growth in both physical capital investment and R&D investment over time. Growth in R&D investment has clearly and significantly outpaced growth in physical investment, a trend that holds across industries and firm size. For example, manufacturing firms alone now spend over $100 billion in R&D annually, more than the (inflation-adjusted) R&D spending of the entire public and private sectors 40 years ago. Both within industries and across industries, there is significant variation in how much investment firms attribute to R&D.

This paper shows that firms that devote more of their investment to R&D activities earn higher risk-adjusted equity returns. It also rationalizes this finding with a model featuring heterogeneous firms and heterogeneous consumers which generates a risk premium for R&D through a novel product market channel. In examining returns to R&D investments, this paper compares those returns to returns on the most significant other form of investment: investments in physical capital. To this end, the measure of a firm's R&D intensity employed in this analysis will be the fraction of total investment expense in a given year allocated to R&D. This metric is clearly an important determinant for firm equity returns: Figure 2

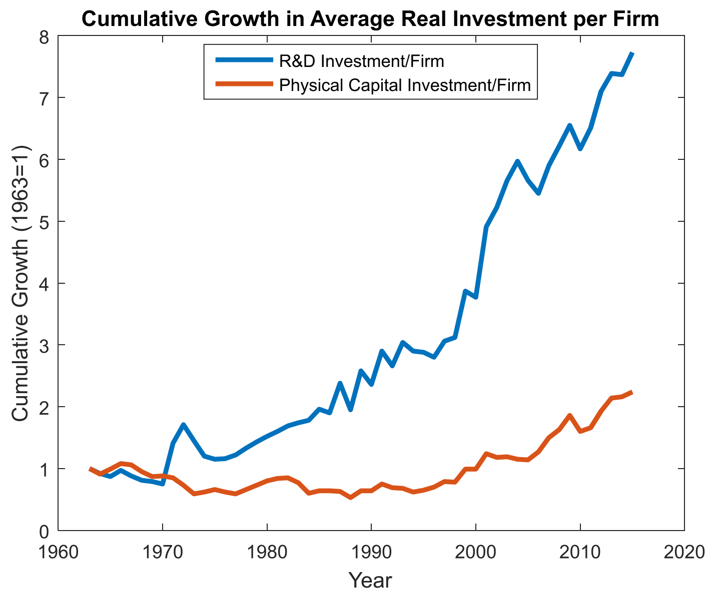

Figure 1: Cumulative Growth in Average Investment per Firm

plots the cumulative annual stock returns for firms with low levels of R&D/total investment, firms with high levels of R&D/total investment, and the aggregate value-weighted market. Firms with a higher ratio of R&D/investment earn significantly higher cumulative returns.

Are these returns compensation for risk? This paper shows that standard risk factors do not explain these returns. Relative to the most common models for expected returns in the literature, firms that allocate most of their investment towards R&D continue to earn higher returns than predicted. Specifically, these firms generate an annual return over 7% higher than that predicted by the Fama-French 3-factor model and an annual return that is 10% above the expected return predicted by the Fama-French 5-factor model. In any rational asset pricing framework, these excess returns must be attributable to some risk factor not spanned by these existing models. Moreover, this paper finds that the risk factor captured by the R&D/investment ratio is important not only to understand the returns of R&D-intensive firms, but also to understand the entire cross-section of stock returns. In Fama-MacBeth tests, the risk factor embedded in these high R&D-intensity firms has a positive and significant price of risk for the entire cross-section.

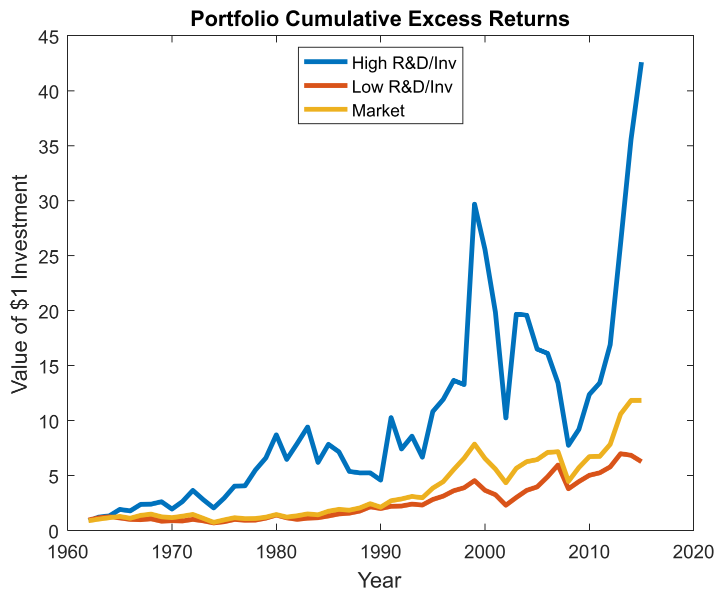

Figure 2: Portfolio Cumulative Excess Returns to $1 invested in January 1962

To understand the risks of these high-R&D firms, it is important to first understand why firms would want to devote a significant fraction of their investment towards intangible capital. While expenditures such as research salaries, blueprint and patent creation, and technology development may not directly enable a firm to produce a higher quantity of products, they often enable it to produce products of a higher quality. Indeed, for many firms, significant R&D expenses are necessary to maintain a competitive market position. This is reflected in the empirical finding documented in the paper that these high-R&D firms also charge significantly higher price markups over cost for their products. Furthermore, the returns on these high-R&D firms can be linked to consumers' expenditures on luxury goods, which are also typically high-markup products (as studied by Ait-Sahalia et al. (2004)).

The final contribution of the paper is then to integrate this key insight that high R&D is connected with higher product quality and price markup into a production-based asset pricing model. The model features heterogeneous firms that create products of different quality levels, with firms that produce higher quality products requiring a greater amount of R&D per unit of physical capital used for production. These firms offer price and quality combinations to heterogeneous consumers, who choose the product that maximizes their utility every period. The economy is subject to total factor productivity shocks which affect the productivity of physical capital and demand shocks which affect the preferences of consumers. The model parsimoniously captures the empirical observation that higher-R&D firms are more exposed to demand fluctuations. This higher exposure explains their higher returns. The model also matches the size and markup dynamics in the data.

This paper is related to several strands of the existing literature. First, there have been several empirical asset pricing studies that focus on how some form of intangible investment affects future equity returns. Perhaps the most noted of these is the 2001 paper by Chan, Lakonishok, and Sougiannis showing that a firm's ratio of R&D expenditures to market equity is related to its future abnormal equity return. Li (2011) studies the interaction between measures of R&D expenditures and financial constraints. Several other papers have studied variants on the R&D anomaly by focusing on firms that have been successful in past R&D (Cohen et al. (2013)) and innovative efficiency (Hirshleifer et al. (2013). These papers have found interesting results in focusing on segments of high R&D-intensity firms that seem to drive the broader set of results. This paper, however, will take the view that another measure of R&D intensity is more informative altogether. In particular, the metric in this paper is important for understanding the returns of larger and more economically important firms and helps to price the entire cross-section of stock returns, two signficiant deviations from the existing literature. Finally, a recent paper by Kogan et al. (2016) has argued that patent creation is important for aggregate economic growth, which will be outside the scope of the analysis considered in this paper.

Second, there have been several structural asset pricing papers which have considered the relationship between equity returns and intangible investment or capital. Lin (2012) is the closest to this analysis; he proposes a model in which intangible capital reduces adjustment costs to physical capital and so drives excess returns of high R&D/Investment firms. Other similar papers include Corhay, Kung, and Schmid (2015), who propose a model linking markups, competition, and stock returns, and Eisfeldt and Papanikolaou (2013), who link organizational capital and expected stock returns. This paper will differ from all of those studies by introducing a different and largely novel mechanism of product market competition to drive the cross-sectional implications. It differs from the first two by also focusing on the cross-section of expected risk-adjusted returns. The R&D/investment measure considered in this paper is also not closely related to the organizational capital measure of Eisfeldt and Papanikolau (2015). Lastly, several papers have focused on the idea that R&D expenditures by firms can drive endogenous growth in the aggregate economy, such as Kung and Schmid (2015) and Ai, Croce, and Li (2012), but this paper will focus on a different channel and will have significantly different cross-sectional implications.

The rest of the paper proceeds as follows. Section 2 describes the data used in the paper, the empirical asset pricing results of the paper, and the empirical motivation for the model. Next, section 3 introduces the model and discusses each component. Section 4 presents the calibration and results of the model. Finally, section 5 concludes.

1.2. **Data and Empirical Results**

1.2.1. Data Sources and Definitions

The data from this project come primarily from the CRSP/Compustat Merged dataset on WRDS. Firm-level accounting data come from Compustat and include balance sheet items (stocks of capital, assets, and capital structure measures) and income and cash flow statement items (revenue, cost of goods sold, R&D expenses, and capital expenditures). Due to the often significant seasonality of many of these series, such as investment expenditures and revenues, observations are collected on an annual basis. The SIC codes 0-999 (agriculture, fishing, hunting), 4900-4999 (utilities), 6000-6999 (financials), 8888 (foreign governments), and 9000-9999 (international/non-operating) were also eliminated for most of the analyses in this paper. Finally, companies that did not report R&D expenditures (about half of the firm-year observations in the sample) were eliminated. The asset pricing results for these firms were quantitatively similar to those for firms that reported zero values of R&D expenditures.

The key ratio of interest for much of this paper will be the fraction of total investment spent on R&D expenses. For this purpose, total investment will be defined as the sum of R&D expenses and capital expenditures. The reporting standard for firms are set by the Generally Accepted Accounting Principles (GAAP), which defines R&D expenses as part of the "planned search or critical investigation aimed at discovery of new knowledge...in developing a new product or service or a new process or technique" or part of the "translation of research findings into a plan or design for a new product or process." R&D is typically differentiated into product development (developing new products or services) and process development (developing new techniques or methods to produce existing offerings) and includes expenses such as research wages, patent development, and software development, with wages making up up to 50% of total R&D in some industries (Hall and Lerner (2010)). Capital expenditures, meanwhile, include all costs in purchasing and making ready for use property, plant, and equipment additions. These typically include long-lived tangible assets such as land, buildings, machinery, equipment, fixtures, and others.

Some have also suggested that selling, general, and administrative (SG&A) expenses and R&D expenses are cross-reported by firms, with some firms reporting as an R&D expense what other firms would report as an SG&A expense (see e.g. Peters and Taylor (2016)). To take this consideration into account, the empirical asset pricing results are also computed using the ratio of the sum of R&D and SG&A expenses to the sum of R&D expenses, SG&A expenses, and capital expenditures. The results are quantitatively similar. Of the firm-year observations for which R&D expenditures are reported, 17% of those observations have reported values of 0 R&D expenses. Of the remaining 83%, the distribution is fairly uniform, as illustrated in Figure 5.

Another definition employed by the paper is a metric of markups, which is used for one result in Section 2.4. While there are many measures of aggregate markups that have been used in similar papers and in the industrial organization literature, in order to get a firm-specific measure of price markups, this paper simply uses the ratio of Revenues to Cost of Goods Sold (COGS) minus 1. This gives the percentage markup over the cost of goods that firms are charging. Others have suggested including other expenses in COGS, such as SG&A expenses, and these changes do not impact the results in Section 2.4. Summary statistics of firms by their level of R&D/Investment are available in Table 4. Higher R&D/Investment firms tend to be smaller, less levered, and have fewer tangible assets. They also tend to charge higher prices relative to product costs and earn higher revenues relative to physical capital (but not total assets, perhaps a reflection of other intangible capital that is included in their asset base).

The Compustat data was then merged with equity returns from CRSP based on the permanent "permco" link between the two. Monthly CRSP returns were collected and are the basis for all of the asset pricing results to follow. From French's website, the monthly excess returns of the market, HML, SMB, RMW, and CMA factors were obtained. Finally, the data series on luxury good sales used by Ait-Sahalia et al. (2004) was also obtained from Yogo's website.

1.2.2. Alpha Sorts

This section is devoted to the analysis of the question of whether the higher returns associated with more R&D-intensive firms are compensation for a recognized risk factor. To analyze this, this paper starts by analyzing the returns of R&D-intensive firms relative to the benchmark model for excess return used in the vast majority of the empirical asset pricing literature, the Fama-French 3-factor model. This model says that excess returns on a given security should be given by equation (1):

$$r_{it}^e = \alpha_i + \beta_{it}^{rmrf} RMRF_t + \beta_{it}^{hml} HML_t + \beta_{it}^{smb} SMB_t + \epsilon_{it} \tag{1.1}$$

where r^e_{it} represents the return on stock i at time t in excess of the risk free rate. If the model holds, then the average α over the cross-section of stocks (or for any portfolio of stocks) should be zero, and the only factors affecting excess returns should be exposures to the excess return on the value-weighted market factor $RMRF$, the high minus low book equity to market equity factor (value minus growth) HML, and the small minus big market cap factor SMB (Fama and French (1992)).

To test this model, firms are sorted into eleven portfolios based on their R&D/Investment ratio. One portfolio contains the firms which report 0 R&D values, while the remaining ten represent the firms which fall into each decile of the R&D/Investment distribution. As Figure 5 illustrates, the approximately uniform distribution of the R&D/Investment measure means that the unit support of R&D/Investment will be fairly evenly divided among these deciles. The portfolios are rebalanced every year with new accounting data, which means that firms will move across deciles as their level of R&D/Investment changes every year. Table 5 gives the transition matrix from one year to the next for the R&D/Investment deciles. The persistence of this measure is significantly higher than those for commonly cited asset pricing anomalies, such as the book-to-market, profitability, investment, and momentum factors, as evidenced by Opp and van Binsbergen (2016).

The value-weighted returns of the firms in the portfolio form the time series of returns for each portfolio. To test the Fama-French 3-factor model, the time series of returns for each portfolio is regressed on the time series of returns for the market excess return and the HML and SMB factors. If the Fama-French 3-factor model correctly prices these assets, then the average excess returns of these portfolio should be explained by their exposures to these three factors and there should be no significant intercept term in the regression results. Table 1 presents the results for the intercept and coefficient terms for each portfolio. The columns of the table represent the various portfolios, from 0 (the portfolio of firms that report R&D values of 0) to the deciles of R&D/Investment (1-10), to “10-1”, which represents the zero-cost portfolio of buying portfolio 10 and short-selling portfolio 1 (similar results hold if one short-sells portfolio 0). Examining the first two rows, which report the intercept (α) regression results and their Newey-West t-statistics, one sees that the pricing errors increase in the R&D/Investment decile, becoming positive and significant for deciles 6, 7, 8, 10 and the long-short 10-1 portfolio. This implies that if one buys portfolio 10 and short-sells portfolio 1, he earns a monthly return of 58bps (7.2% annualized) in excess of what the Fama-French 3-factor model would predict based on this portfolio's exposures to the three Fama-French factors. This α is statistically significant at the 5% level and thus indicates a violation of the model.

	R&D/Investment Decile											
	0	1	2	3	4	5	6	7	8	9	10	10-1
α	-0.026	-0.057	0.028	0.065	-0.026	0.094	0.219	0.301	0.470	0.099	0.523	0.580
	(-0.36)	(-0.61)	(0.32)	(0.75)	(-0.30)	(1.01)	(2.33)	(2.63)	(3.34)	(0.61)	(2.90)	(2.78)
$RMRF$	0.909	0.890	1.064	1.046	1.037	1.015	0.950	1.023	1.036	1.025	1.095	0.206
HML	0.130	0.155	0.153	-0.036	-0.139	-0.299	-0.354	-0.450	-0.602	-0.830	-0.794	-0.950
SMB	-0.142	-0.226	-0.158	0.078	0.004	0.035	0.022	0.137	0.238	0.592	0.848	1.074

Table 1: R&D/Investment Decile Portfolio Regressions on Fama-French 3 Factors. First row gives Fama-French 3-Factor alphas, with Newey-West t-statistics below. Bottom 3 rows give portfolio betas with respect to Fama-French 3 factors.

The bottom three rows also reveal information about the risks of the firms in these portfolios. Exposure to the aggregate market risk factor, RMRF, is roughly constant across deciles, indicating that there are no significant differences in the exposures of the high-R&D versus the low-R&D firms to the risks spanned by the aggregate stock market. This is an interesting result in itself as it suggests that the connection between R&D-intensive firms and high-beta firms is not as close as some have suggested, which may be due to the focus on just the composition of investment (rather than the composition and amount of investment which some measures combine). There are, however, significant differences in the exposures of these firms to HML and SMB. High R&D/Investment firms tend to load far more negatively on HML than do low R&D/Investment firms. This pattern indicates that these firms have lower ratios of book equity to market equity and are more growth firms than value firms, consistent with what one might expect. Similarly, high R&D/Investment firms load more positively on SMB than do low R&D/Investment firms, indicating that these firms tend to be smaller. Still, after accounting for these exposures (which have somewhat offsetting effects due to their opposite signs), the Fama-French 3-factor model fails to fully explain the higher returns of the high R&D/Investment firms.

In response to the documented failures of the Fama-French 3-factor model to explain certain anomalies related to investment and profitability (see e.g. Hou et al. (2015) and Novy-Marx (2013)), Fama and French updated their model to include two additional factors designed to capture the variation in expected returns associated with firms' investment policies and profitability. The updated model is expressed in equation (2):

$$r_{it}^{e} = \alpha_i + \beta_{it}^{rmrf} RMRF_t + \beta_{it}^{hml} HML_t + \beta_{it}^{smb} SMB_t + \beta_{it}^{rmw} RMW_t + \beta_{it}^{cma} CMA_t + \epsilon_{it} \quad (1.2)$$

where the first four terms on the right hand side are identical to the previous model and the excess return now includes compensation for the return's exposure to RMW, which represents robust operating profitability minus weak and CMA, which represents conservative investment minus aggressive. Note that operating profitability is computed as $\frac{Revenues - COGS - Interest - SG\&A}{Book\,Equity}$ and so does not explicitly include either R&D expenses or capital expenditures. Similarly, Fama and French define investment as the growth of total assets in the previous year divided by the amount of assets two years past, so neither measure is mechanically linked to the R&D/Investment measure in this paper (Fama and French (2015)).

Given that these revisions explicitly seek to address investment-based anomalies, it is natural to ask whether the R&D/Investment factor which targets the composition of investment can still be used to generate portfolios which earn abnormal returns under the Fama-French

5-factor model. One can perform the same test as before: run a time series regression of the value-weighted returns of each R&D/Investment portfolio on the five Fama-French factors and report the intercept and coefficient values for each portfolio, checking to see whether the portfolios generate positive intercepts. Table 2 gives those results.

Beginning again with the regression intercepts, an even more striking pattern is apparent. Not only are the α values still increasing in the R&D/Investment deciles, but the magnitudes are now significantly higher. Whereas the long high R&D/Investment and short low R&D/Investment portfolio earned risk-adjusted monthly excess returns of 58bps per month under the Fama-French 3-factor, it now earns risk-adjusted excess returns of 82bps per month, or over 10% per year, relative to the Fama-French 5-factor model. One can see why this is the case by examining the factor exposures of this portfolio. The first three coefficients follow similar patterns to their counterparts in the Fama-French 3-factor results. Namely, exposures do the market do not seem to have a significant pattern as R&D/Investment varies but high R&D/Investment firms seem to be more growth firms (negatively exposed to HML) and smaller firms (positively exposed to SMB).

What is different is that these firms are also fairly negatively exposed to the profitability factor, RMW. This result is somewhat consistent with the observation that these firms also tend to have lower Revenue/Asset ratios and lower Asset/Book Equity ratios. Interestingly, the high-R&D/Investment firms and low R&D/Investment firms have very similar exposures to the investment factor. This latter observation suggests that the decomposition between the amount of investment and the composition of investment is an important one. Combined, the exposures to RMW and CMA lower the benchmark for the returns of these high R&D/Investment portfolios under the Fama-French 5-factor model and thus lead to the more significant intercept term.

Tables 21-23 demonstrate that this measure is also robust to other firm-level characteristics.

	R&D/Investment Deciles											
	0	1	2	3	4	5	6	7	8	9	10	10-1
α	-0.077	-0.054	-0.015	0.085	0.023	0.206	0.240	0.361	0.769	0.325	0.768	0.821
	(-1.01)	(-0.53)	(-0.16)	(0.90)	(0.27)	(2.17)	(2.35)	(3.06)	(5.39)	(2.00)	(4.32)	(3.18)
$RMRF$	0.929	0.900	1.076	1.035	1.028	0.997	0.946	1.005	0.959	0.993	1.070	0.169
HML	0.066	0.091	0.106	0.003	-0.076	-0.287	-0.403	-0.423	-0.410	-0.875	-0.860	-0.951
SMB	-0.130	-0.254	-0.136	0.077	-0.007	-0.035	0.001	0.093	0.106	0.406	0.632	0.887
RMW	0.044	-0.109	0.103	0.010	-0.075	-0.252	-0.056	-0.179	-0.523	-0.706	-0.868	-0.758
CMA	0.152	0.146	0.107	-0.075	-0.125	-0.050	0.092	-0.091	-0.476	0.017	0.095	-0.050

Table 2: R&D/Investment Decile Portfolio Regressions on Fama-French 5 Factors. First row gives Fama-French 5-Factor alphas, with Newey-West t-statistics below. Bottom 5 rows give portfolio betas with respect to Fama-French 5 factors.

Higher R&D/Investment is positively associated with higher risk-adjusted equity returns after controlling for both gross and net profitability. Within profitability quintiles, firms which do more R&D relative to total investment earn significantly higher Fama-French 3-factor and 5-factor alphas. Similarly, double-sorting first by the amount of investment (relative to total assets) that a firm spends and then by R&D/Investment still produces increasing alphas in the R&D/Investment ratio. This effect is present for all but the firms which invest the least (those in the bottom 20% of the Investment/Assets ratio.) These results are also robust to other factor-based models for equity returns. Table 24 demonstrates that this effect is amplified when the Quality-minus-Junk factor proposed by Asness et al. (2013) is included as a risk factor. Similarly, Table 25 presents the robustness of these results to the inclusion of a momentum factor.

There are several important differences between these results and those of earlier papers. First, portfolio returns are value-weighted rather than the equal-weighted returns found in the literature. Value-weighting these returns is important for several reasons. First, value-weighting portfolio returns both prevents big price changes to small market cap firms from having an outsized impact on portfolio returns. Thus value-weighting focuses on the larger and more central firms and is thus more meaningful from an economic standpoint. Table 15 shows this explicitly by double-sorting firms into groups based on assets and then R&D measures. For the R&D/Investment metric employed in this paper, there is a significant return differential attributable to R&D-intensive firms across all size quintiles. In comparison, for the R&D/Market Equity measure, the effect is only significant at a 5% level or stronger for the smallest quintile of firms, whose median asset values are 0.25% of those of firms in the highest quintile. This pattern is similar for most other existing R&D measures in the literature.

Second, value-weighting is more in keeping with the asset allocation of one who would hold this portfolio. Equal-weighting requires constant rebalancing of a portfolio and very high associated transactions costs. It is difficult to interpret the returns from such a portfolio as an asset pricing anomaly if the costs of exploiting the strategy outweigh the potential benefits. Thus it is important to value-weight the returns within a portfolio. It is even more important when one considers that the significant asset pricing results that one obtains using equal-weighted portfolio returns with existing asset pricing measures disappear under value-weighting of portfolio returns, as shown in Table 6. Thus, the equal-weighted results arbitrarily overemphasize the importance of small market cap firms and may not truly represent an anomaly. By showing that the results in this paper hold for value-weighted (as well as equal-weighted in Table 7) these concerns are eliminated. Another important difference the measures of R&D intensity are fairly different. Tables 16-20 report the similarities between the measure introduced in this paper and five other common measures: R&D/Market Equity, R&D/Sales, R&D Capital/Market Equity, R&D/R&D Capital, and SG&A Capital/Assets. The percentage of firms which fall in the same decile when sorted by the measure in this paper as when sorted by these measures is fairly low. This is also evidenced by the fact that the measures themselves are not highly correlated: the highest Pearson or rank correlation between the measure in this paper and any of the existing measures is 0.4. These low correlations reflect an important economic insight captured by the R&D/Investment measure: its focus on the composition of investment. While other measures combine the composition of investment and the amount of investment, the focus on what type of investment a firm is doing, rather than how much it is doing, is an important differentiating factor of this measure. Finally, this measure is the first R&D-based measure to be significantly priced in the full cross-section. This is important as it means that the empirical asset pricing results in this paper are important for understanding the entire cross-section of equity returns.

The Fama-French 3-factor results and 5-factor results clearly indicate that high R&D/Investment portfolios earn significant positive risk-adjusted returns. These firms tend to be smaller, more growth firms, and less profitable by the Fama-French metric but to load similarly on the aggregate market factor and the investment factor as their low R&D/Investment counterparts. Despite these differential loadings, the higher returns of the high R&D/Investment portfolios are not rationalizable by any of the Fama-French factors.

1.2.3. Fama-MacBeth Results

In any rational asset pricing model, these higher returns must be attributable to some additional source of risk faced by these high R&D/Investment firms and affecting investors' discount rates. A natural follow-up question would then be to ask whether the risk encapsulated in these high R&D/Investment firms is important only for them or for a broader number of firms. This risk is captured by the return of the portfolio going long high R&D/Investment firms and short low R&D/Investment firms, so, to answer this question, this section presents the cross-sectional asset pricing results using this "10-1" portfolio as a factor.

A Fama-Macbeth procedure tests whether this innovation risk factor measure is priced for a larger cross-section of assets. For this analysis, both industry portfolios and the entire dataset of monthly stock returns (not just those of stocks who report R&D values). For the industry portfolio results, the largest cross-section of industries categorized by Fama and French is used for the longest timespan for which all data is available. This results in 49 industries over a period of 559 months. The Fama-Macbeth procedure is computed as follows. First, for each industry/stock at each point in time, rolling-window betas with respect to the High R&D/Investment minus Low R&D/Investment (henceforth referred to as the innovation factor or R&D) and either the Fama-French 3-factor model or the Fama-French 5-factor model. To account for possible covariances between the factors, these betas are estimated simultaneously in two groups: one group with the innovation factor and the Fama-French 3 factors and one group with the innovation factor and the Fama-French 5 factors, given by equations (1) and (2), respectively. After that, at each point in time, a cross-sectional regression of excess returns on betas is computed (again, separate regressions for the 3 factors and R&D from the 5 factors and R&D), and the prices of risk extracted as the λ values in specifications (3) and (4):

$$r_{it}^{e} = \alpha_i + \beta_{it}^{rmrf}\lambda_t^{rmrf} + \beta_{it}^{hml}\lambda_t^{hml} + \beta_{it}^{smb}\lambda_t^{smb} + \beta_{it}^{R\&D}\lambda_t^{R\&D} + \epsilon_{it} \tag{1.3}$$

$$r_{it}^{e} = \alpha_i + \beta_{it}^{rmrf}\lambda_t^{rmrf} + \beta_{it}^{hml}\lambda_t^{hml} + \beta_{it}^{smb}\lambda_t^{smb} + \beta_{it}^{rmw}\lambda_t^{rmw} + \beta_{it}^{cma}\lambda_t^{cma} + \beta_{it}^{R\&D}\lambda_t^{R\&D} + \epsilon_{it} \tag{1.4}$$

The choice of test assets (whether industries or individual stocks) represents a trade-off between accuracy of beta estimation and use of a larger cross-section. While the use of portfolios mitigates the errors-in-variables problem associated with estimating time-varying betas for individual stocks and using those estimated betas in a second-stage estimation, the construction and number of the portfolios need to be carefully considered. Industry portfolios are utilized here so that the potential issue of sorting portfolios by characteristics is avoided. Moreover, the broadest set of industries reported is used to obtain as large a cross-section (and hence as powerful a test as possible.) The procedure is also replicated with individual stocks and the results are displayed in Table 9 in the appendix.

Table 3 presents the Fama-Macbeth results for the Fama-French 3-factors and the innovation factor, using the 49 industry portfolios as defined on French's website. Both equally-weighted and value-weighted portfolio returns are considered (with the factor weighting matching the return weighting); the results are robust to choice of portfolio and factor weighting. Across the specifications, both the market factor and the innovation factor both have positive and significant coefficients. Thus, exposure to the market and innovation risk factors are associated with significantly higher expected returns. In particular, an increase by one in market beta is associated with an increased monthly industry return of 42-59 bps, and an increase by one in innovation beta is associated with an increased monthly industry return of 86-89 bps. Moreover, SMB is not significant in either specification and HML is only significant in one of the specifications. Thus, these results suggest that both innovation risk and market risk have significant pricing power for the cross section of industry returns.

These results are robust to other specifications. Table 8 in the appendix presents the industry Fama-Macbeth results for the innovation factor and Fama-French 5 factors. In those tests, the market risk factor is the most significant, followed by the innovation risk factor (which is significant for equally-weighted portfolio returns but less so for value-weighted portfolio returns). None of the other factors are significant in either test. Table 9 in the appendix presents the Fama-Macbeth results for value-weighted individual stock returns. With respect to the three-factor model, the innovation factor is significant at the 10% level and has a t-statistic fairly similar in magnitude to that of two other factors, the excess market return and SMB. HML, however, is no longer significant with either specification. More formally, the test of whether HML is spanned by the RMRF, SMB, and innovation factors produces no significant intercept for HML and thus indicates that it is spanned by the other three. In contrast, the innovation factor is not spanned by the three traditional Fama-French factors. For the five factor results, the innovation factor is significant at the 5% level. The same patterns from the 3-factor model are still apparent, namely that RMRF and SMB are significant but HML is not. Additionally, CMA is no longer significant (again indicating that this contrast between the composition of investment and the intensity of

	λ Values	
$High - Low$ $\frac{RD}{Inv}$	0.888* (1.83)	0.856** (2.35)
$Mkt - rf$	0.591*** (2.89)	0.425* (1.77)
HML	0.156 (0.97)	0.408** (2.15)
SMB	0.078 (0.48)	0.249 (1.27)
Weights	Value	Equal

Table 3: Fama-Macbeth Industry Results for Fama-French 3 Factors and Innovation Factor

This table reports the Fama-Macbeth prices of risk from the two-stage Fama-Macbeth regression presented in the paper for value- and equal-weighted industry returns. Values reported as percentage points per month. First, for each industry at each point in time, 72-month rolling-window betas with respect to the innovation factor and simultaneously the Fama-French 3-factor model. After that, at each point in time, a cross-sectional regression of excess returns on betas is computed, and the prices of risk extracted as the λ values in specification (3). Specifically, each estimate reported is an estimate of a lambda value based on a time-series average of the lambdas estimated for each cross-sectional second-stage regression. Results robust to value vs. equal weighting of factors and different horizons for rolling window estimation. *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. Numbers in parentheses are t-statistics corrected for autocorrelation over time.

investment is important) and RMW has a negative price of risk.

Across all of these specifications, the key result is the consistent significance of the innovation factor, which indicates that the risk spanned by this factor is important for the entire cross-section of excess industry and stock returns, and that it is not spanned by the other factors in the Fama-French models. Second, adding the innovation factor seems to eliminate the explanatory power of HML for the cross-section of equity returns.

1.2.4. Evidence on Markups and Luxury Goods

The previous sections have clearly documented that high R&D/Investment firms earn not only higher equity returns, but also higher returns after accounting for the most common equity return factor model predictions. Rational asset pricing models imply that this must be a compensation for some form of risk which matters to investors, and the Fama-MacBeth results indicate that this risk matters not only for the returns of high R&D/Investment firms, but also for the entire cross-section of equity returns. Having documented this return pattern and the importance of the risk factor spanned by these high R&D/Investment firms, the goal is now to try to understand this risk factor and to link it to underlying economic quantities.

For this task it is helpful to think about why these firms do so much R&D relative to other forms of investment. Numerous theories for this have been advanced, but one that has received general acceptance is that R&D acts as a way for firms to maintain their competitive advantages. In particular, for firms that produce differentiated products and rely on being able to charge premiums for those products, one important way to maintain their differentiation and the willingness of consumers to pay premia is to continue to innovate. Indeed, the summary statistics confirm that firms which do more R&D relative to total investment charge higher price markups relative to cost (have a higher Revenue/Cost of Goods Sold ratio.) The results in Table 10 provide further stronger evidence on this point–the regression results of future markup on the ratio of R&D/Investment, firm size, and controls for industry and year fixed effects indicate a positive and significant effect of R&D/Investment on future firm markups. This evidence is also consistent with firm-level evidence by Cassiman and Vanormelingen (2013), who find that product innovations increase firm markups by an average of 5.1% and process innovations increase firm markups by 3.8% on average.

But why are higher markups in themselves more risky and why is this risk important to investors? One clue comes in the relation between the returns to the long-short R&D/Investment

portfolio and the luxury sales index compiled by Ait-Sahalia et al. (2004). Table 11 shows that the growth in luxury good sales helps to explain the risk premium associated with this long-short portfolio, which indicates that the risk implied by the sale of luxury goods, another high markup item, is similar to that spanned by the R&D/Investment factor. In particular, the luxury consumption risk that Ait-Sahalia and his coauthors identify seems to be linked to the risk of these high R&D/Investment firms. One can think of this more broadly as demand risks that affect all firms, but particularly firms which rely on being able to charge high markups, as both their prices and quantities are potentially more prone to shifts in consumer preferences. The next section builds on this key insight and introduces a model which formalizes this intuition.

1.3. **Model Setup**

The model features a number of different elements which will be described in this section. Its most novel feature is the integration of product market competition into a production-based asset pricing framework, as will be described later. The model is infinite horizon and discrete time and contains heterogeneous consumers, heterogeneous firms, and two state variables. The consumers, firms, and environment are described in the following subsections.

1.3.1. Consumers

Consumers in this model are the source of demand for the firms. Each period, a unit mass of consumers enters the market for a good and views the menu of options offered by firms. An option offered by a firm consists of a quality level of product and corresponding price, as will be discussed in greater detail later. For now, it suffices to say that quality is a feature which vertically differentiates products–that is, all consumers prefer a higher quality product, all else equal. Each consumer evaluates the menu of offerings and chooses the product that maximizes his utility. The consumer may choose to buy either one unit of one product or not to buy at all. What differentiates consumers is their willingness to pay for an increase in a product's quality.

Formally, the willingness of consumer j to pay for a higher quality product is represented by the parameter θ_j. The quality of products is indexed by s and the indirect utility that a consumer with preference parameter θ_j maximizes every period is given by:

$$U_j = \left\{ \begin{array}{cc} u_0 + \theta_j^s - p & \text{if purchase good of type s at price p} \\ 0 & \text{else} \end{array} \right\}$$

That is, consumers get some base utility from purchasing a product of any quality level u_0, then some utility which depends on both their preference for quality and the quality of the good they purchase. Finally, they internalize the price of the good which they purchase. This framework is a fairly standard one for vertically differentiated goods, see e.g. Tirole (1988). The consumers are price takers and so have no ability to impact prices; their decisions are independent of the decisions of the other consumers. Therefore the consumer's problem can be expressed as:

$$\underset{s_i}{max} \left\{ u_0 + \theta^{s_i} - p\left(s_i\right), 0 \right\} \tag{1.5}$$

The parameter θ is the only differentiating factor among consumers and thus what drives any differences in their decisions. In wanting to tie the distribution of θ to empirical counterparts, several options were considered. Existing evidence suggests that factors which derive heterogeneity across households include income, age, wealth, and other sources, but, for the purposes of this model, perhaps income is the most salient. The distribution of income has been extensively studied and researchers have suggested a number of different distributions to match the cross-sectional patterns of income, including the exponential distribution, the lognormal distribution, and the generalized Pareto distribution. Among these, the exponential is chosen in this paper because of its property as a distribution governed by one parameter. Given that there is no precise data counterpart to the preference parameter, the goal is to calibrate as close to the data as possible, and having relatively fewer parameters for the distribution of preferences helps achieve that goal. The results, however, are robust to the other distributions with similar properties.

The dynamics of the θ distribution vary over time in the model with one of the model's two state variables, X_t. This state variable can be interpreted as a demand or preference shock and it is set equal to the mean (or the inverse of the scale parameter) of the exponential distribution of θ. That is, higher values of X_t imply a distribution which skews more towards greater willingness to pay for quality, while lower values imply a distribution which skews more towards lower willingness to pay for quality, as illustrated in Figure 6. The log of X_t will follow an AR(1) process, given by equation (6).

$$x_{t+1} = (1 - \rho_x)\,\bar{x} + \rho_x x_t + \sigma_x \epsilon^x_{t+1} \tag{1.6}$$

where ϵ^x_{t+1} is a standard normal random variable. The parameters for this process can be directly tied down by the aggregate markups in the model economy, as will be discussed in the calibration section.

1.3.2. Firms and Good Quality

The other type of agent in the economy is a firm. Firms will produce the goods and will face constraints on their production from both consumer demand and the supply and productivity of capital. Firms will maintain two types of capital–intangible and physical–which will affect their profits and values differentially. Finally, firms will be heterogeneous, with a sufficient statistic for firm heterogeneity being the quality level of products that they produce. Firms will choose both their quality level and the price which they set for their goods, and this will determine their capital needs and profit.

At any given point in time, firms maintain two stocks of capital. One is physical capital

K_t, which is required for production. Production follows a standard AK-technology, where total output Y_t can be represented as $Y_t = A_t K_t$ and A_t represents the state of capital productivity in the economy. A_t is the second state variable in this model and again its log follows an AR(1) process, displayed in equation (7).

$$a_{t+1} = (1 - \rho_a)\,\bar{a} + \rho_a a_t + \sigma_a \epsilon^a_{t+1} \tag{1.7}$$

where ϵ^a_{t+1} is a standard normal random variable uncorrelated with ϵ^x_{t+1}. Firms also maintain levels of intangible capital IK_t. Unlike physical capital, intangible capital is not used directly in the production process. Rather, it helps to differentiate products. Specifically, firms that maintain higher levels of intangible capital relative to their total capital produce higher quality products. For a firm i with capital levels K_{it} and IK_{it}, the quality of goods that it produces is $s_{it} = \frac{IK_{it}}{IK_{it}+K_{it}}$. The intuition for this feature is as follows: physical capital is needed for the actual production of products, but intangible capital helps to differentiate goods. The more intangible capital, in the form of research, thought, innovation, testing, process development, etc. that goes into the product, the higher the quality that the product will be and the more that consumers will be willing to pay for it. This will also help match the empirical finding that firms that do more R&D/Investment charge higher markups over cost for their products. One could akternatively envision N production functions for goods of different qualities requiring certain ratios of physical and intangible capital. Such a setting would be isomorphic to this one, and the flexibility of the setup allows for a wide range of interpretation of R&D expenses.

There are a finite number of quality levels s_1, s_2, $\ldots$ s_N at which a product can produced, evenly spaced throughout the unit interval that defines the quality spectrum. The economy features N firms; at time 0, firm i is born into quality level s_i. That is, at time 0, there is exactly one firm per quality level. As long as this continues to be the case, this single firm acts a monopolist in the market for goods of that specific quality level. Of course, the prices that such a firm can set will be constrained by the prices set by other firms, but this firm can earn positive profits. If multiple firms are producing the same quality product, however, these firms will engage in Bertrand competition and their profits will both be 0. While there is no firm entry, firms are able to endogenously choose their quality level and switch into any of the N good quality markets. As Proposition 1 states (and Appendix B proves), as long as there is some positive switching cost c, this will never happen in the Pareto efficient Nash equilibrium.[1]

Proposition 1. *If there is a positive cost of switching quality levels, no firm will ever switch from its initial quality level in the Pareto efficient Nash equilibrium.*

Firms rent their physical and intangible capital at rental rates r_K and r_{IK} respectively. One could view this as resulting from the inelastic supply of these two inputs from an unmodeled section of the economy. One could also alternatively think of firms owning their capital and being able to freely adjust it intro-period after observing the aggregate state variables. In that case the costs of capital would be the difference between the price paid today and the discounted depreciated amount for which it can be sold in the following period and so would be stochastic. The quantitative results of the model, however, do not change under this formulation.

1.3.3. Final Goods Market and Firm Problem

Figure 3 illustrates the interactions of consumers and firms in the market for final goods.

Every period, each firm (representing one of the N different quality levels) decides on a price to set for its product after observing the realizations of the two state variables. Based on that price and the prices and qualities set by all of the other firms, consumers will choose the product offering which maximizes their utility. The aggregation of consumers who choose a

[1]This result is similar to that in Chapter 3 of Grossman and Helpman: "Innovation and Growth in the Global Economy".

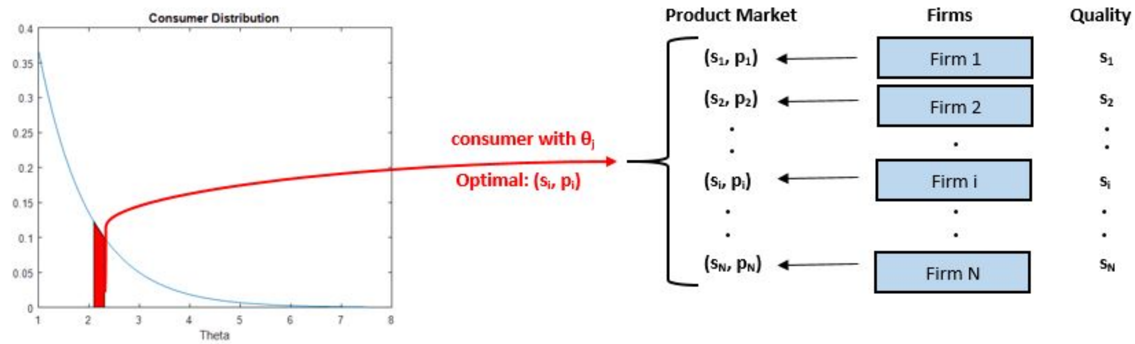

Figure 3: Interaction of Consumers and Firms in Product Market

particular product will determine the quantity demanded for that product. In equilibrium, firms will know this quantity, and so will rent the exact amount of physical and intangible required to product that many units of their quality level.

Each firm will set its price taking into account its own quality level and those of other firms, as well as its beliefs about the prices that the other firms set (which will be correct in equilibrium). It will account for the two aggregate states in the economy, A_t and X_t, and the decision-making problem of the consumers. Since capital is adjustable each period, firms will incorporate the cost of capital into their pricing decision, and, by choosing the price they set (and thus the quantity they will sell), they also choose their optimal capital levels. As a result, the solution methodology does not require tracking the capital distribution of firms over time. The firm's problem can thus be written as maximizing profit given in equation (8) subject to the demand constraint in equation (9).

$$\pi_{it}\left(A_t, L_t, s_i\right) = \underset{P_{it}, s_i}{max} \left\{ P_{it} A_t K_{it} - r_K K_{it} - \underbrace{\frac{r_{IK} s_i K_{it}}{(1-s_i)}}_{r_{IK} IK_{it}} - c1_{\{\text{switch}\}} \right\} \tag{1.8}$$

$$\text{s.t. } A_t K_{it} = \int 1_{\{argmax(U_j)=s_i\}} f(j)\, dj \tag{1.9}$$

where the latter conditions enforces that firms produce exactly enough to meet the quantity demanded. (One could also make this last equality a weak inequality and have firms maximize over their capital stocks, but it is clear that no firm would want to rent more capital than required to meet its demand.)

1.3.4. Firm Value

Firms are entirely equity financed and earn profits that are weakly greater than 0 every period. As a result, the value of a firm is simply its discounted dividend stream, where the dividends are equal to the profits earned by a firm in a given period. Firm value V_t can be expressed as:

$$V_{it}(A_t, X_t, s_i) = \pi_{it} + E_t[M_{t+1} V_{it+1}] \tag{1.10}$$

where M_{t+1} represents the stochastic discount factor in the economy. This discount factor has exposures to both shocks A_t and X_t, as well as time-varying prices of risks for both shocks. M_{t+1} is thus given by equations (11) - (13).

$$log(M_{t+1}) = log(\beta) + \gamma_{at}(a_t - a_{t+1}) + \gamma_{xt}(x_t - x_{t+1}) \tag{1.11}$$

$$\gamma_{at} = \gamma_{a0} + \gamma_{a1}(a_t - \bar{a}) \tag{1.12}$$

$$\gamma_{xt} = \gamma_{x0} + \gamma_{x1} (x_t - \bar{x}) \tag{1.13}$$

1.4. Model Calibration and Results

1.4.1. Model Calibration

Despite the novel product market dynamics in the model, the number of parameters to calibrate is quite limited. The model contains 13 parameters and is calibrated on a monthly basis. The parameters can be grouped into four categories: the SDF parameters β, γ_{a0}, γ_{a1}, γ_{x0}, and γ_{x1}, the rental rates r_K and r_{IK}, and the parameters governing the productivity shock $\bar{a}$, σ_a, and ρ_a and those governing the preference shock $\bar{x}$, σ_x, and ρ_x.

Two groups of these parameters can be estimated directly from the data. First, the rental rates are tied to the rates of depreciation on the two forms of capital, as discussed in Section 3. Thus, these monthly parameters can be tied to the annual depreciation rates of physical and intangible capital found in Lin (2012). Lin finds the rate of depreciation on tangible capital to be 0.1 and the rate of depreciation on intangible capital to be 0.2. Given that these rental rates also reflect some cost of discounting, the monthly rental rates for tangible and intangible capital are set at 0.01 and 0.02, respectively. Second, the parameters governing the productivity shock are standard in much of this literature and have been estimated and used by a number of papers. This paper follows the calibration in Zhang (2005) for the monthly AR(1) process governing productivity shocks.

This leaves two sets of parameters which have to be calibrated. The first is the set of three parameters governing the demand shock process. Given the difficulty of observing this process directly in the data, the challenge is to find a readily observable empirical series to which the demand shock can be closely linked. In this model, given the previous two sets of parameters, the demand shock determines the markups set by each firm. Thus, one can directly link the demand shock to the aggregate markup that this model produces. Fortunately, Kung, Schmid, and Corhay (2015) have estimated an AR(1) process for the aggregate price markup series, and so this paper calibrates the demand shock process to most closely match the parameters that they estimate. The values for the parameters in the paper can be found in Table 12 and the resulting values for the aggregate markup process in Table 13.

The last set of parameters is the one governing the SDF. There are five parameters in the SDF, and thus one needs at least five data counterparts with which to identify these parameters. This paper follows the lead of Zhang (2005) in taking three of these data counterparts to be the mean and volatility of the risk-free rate and the Sharpe ratio of the market portfolio. The remaining two parameters are used to target the returns to low R&D/Investment firms and the returns to high R&D/Investment firms. All of the estimated parameters can be found in Table 12, while the risk-free rate and market return moments can be found in Table 13. The returns to high and low R&D/Investment firms will be discussed in the next section.

1.4.2. Model Results

The most novel feature of the model is the product market dynamic and thus it makes sense to start there. While the effects of the productivity shock are fairly standard, the effects of the consumer preference shock are perhaps not so readily understood. One way to understand the effects of these shocks is to look at their impacts on the decisions which consumers make. Figure 4 illustrates the effect of a shock to preference on consumer decisions.

The graph on the left represents the distribution of consumers under the median state of consumer preference, while the graph on the right represents the distribution of consumers following a positive preference shock. The x-axis displays the θ parameter of consumers while the y-axis gives the density (which integrates to 1 in both cases.) Ignoring the

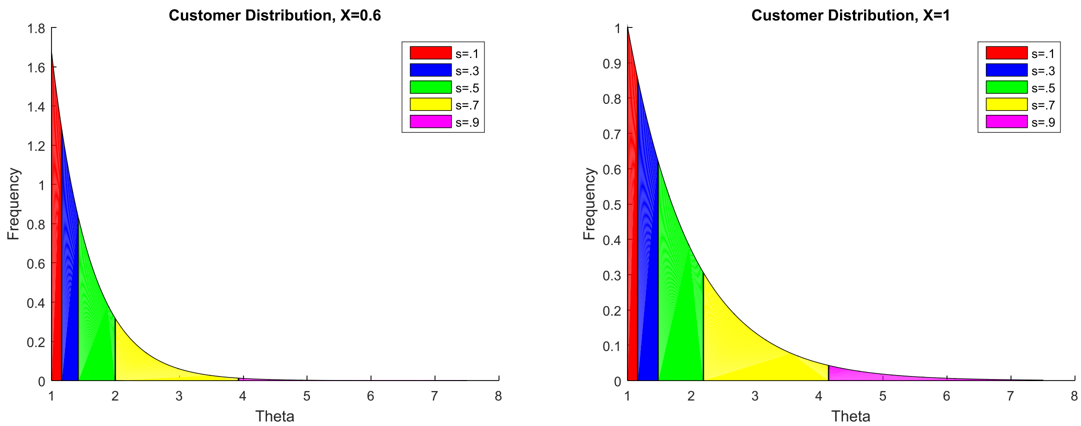

Figure 4: Effect of a Shock to Consumer Preferences on Decisions and Quantities
X represents the aggregate demand state of the economy, while theta represents the preference parameter of a consumer

colors, one notices that the distribution shifts towards higher θ values following a positive preference shock. What is more interesting is the effect that this has on the product choice of consumers. The colors on the graph represent the product choices that consumers make, going from red (lowest quality) to magenta (highest quality). One sees that consumers shift significantly towards higher quality products following a positive preference shock and shift away from products of lower qualities. Figure 7 shows that this example is consistent with the global behavior of the model: the higher the preference parameter, the lower the market shares of low-quality products and the higher the market shares of high quality products. This is in keeping with one's intuition: as consumers' tastes shift more towards higher quality, we should see the market shares of these companies growing.

The other component to the firm's profits, besides its market share, is its profit margin, or the amount of profit that it earns for every unit that it sells. Figure 8 plots the profit margins of low-, medium-, and high-quality firms as a function of the underlying preference state. Consistent with empirical evidence (see e.g. Nekarda and Ramey (2013)), markups are fairly insensitive to demand shocks. However, the markups of high-quality firms are more procyclical with respect to demand shocks than those of low- or medium-quality firms, whose markups are essentially acyclical. Combined with the differential exposures of quantities to demand shocks, this implies that the profits (cash flows) of high-quality firms covary much more positively with demand shocks than do those of low-quality firms, whose cash flows covary somewhat negatively with demand shocks.

This hetereogeneous exposure to demand shocks is key for the asset pricing implications of the paper. Firms of all qualities have similar exposures to productivity shocks. This is because a positive productivity shock reduces the amount of capital required to produce a certain amount of output, which also reduces the amount of intangible capital required. The former effect is more significant for low-quality firms and the latter is more significant for high-quality firms, but, on balance, the total effects are similar. What differentiates firms, then, is their exposure to the demand shock. Since both shocks are priced by investors, the similar loading of all firms to the productivity shock will drive a risk factor which is essentially common to all firms. In a CAPM sense, this will be the risk factor that the market prices. The covariance with respect to the demand shock will be a priced risk factor whose quantity of risk differs significantly across firms, and this is what will drive the heterogeneity in returns. Since the high-quality firms are more exposed to this shock, they will earn higher returns in equilibrium.

This result is consistent with the earlier empirical evidence on the reliance of these firms on markups and the comovement of their returns with the growth of luxury sales. It is precisely these quantities that are most directly tied to preference shocks as those shocks most strongly affect whether consumers purchase high-markup products and are willing to pay high premia for them. It should thus not be surprising that this risk factor also drives the higher returns of R&D-intensive firms that are also dependent on markups.

Table 14 presents the models results on returns and firm size. The model matches the CAPM alpha and beta results fairly well. In addition, despite not being calibrated to match firm sizes, the product market implications for firm size match the empirical distribution of firm size by ratio of R&D/Investment closely. High R&D/Investment firms tend to be smaller and to earn higher CAPM alphas despite their slightly higher CAPM betas. The fact that these firms are smaller on average is important for the model. Since the magnitude of the covariance of these firms' cash flows with the demand shock is much higher than the magnitude of the covariance for the lower R&D/Investment firms, one needs that these firms account for less of the market value than the low R&D/Investment firms in order for the value-weighted market portfolio to have minimal exposure to the demand shocks. Finally, the model does not match the Fama-French 3 or 5 factor models simply because there are not enough shocks in the model to capture asset pricing cross-sectional heterogeneity on more than two dimensions, but this is an interesting area for future research.

1.5. Conclusion

This paper focuses on the returns to this innovation at the firm level by proposing a novel characteristic that examines the fraction of investment attributed to R&D expenses. This metric is important for the entire cross-section of firms, including larger, more economically significant firms. The asset pricing implications examine how risk and returns vary with the composition of investment chosen by a firm.

Relative to the Fama-French 3-factor model, a portfolio which goes long the highest R&D/Investment firms and short the lowest R&D/Investment firms earns monthly excess returns of 58bps per month, or just over 7% annually. This is after accounting for the fact that these high R&D/Investment firms have slightly higher betas and tend to be small, growth firms. These results are significant at the 1% level, and, in contrast to previous studies on R&D, hold for value-weighted portfolios. Compared to the Fama-French 5-factor model, the results are even more significant: the long-short portfolio earns risk-adjusted returns of 82bps per month, which corresponds to over 10% per annum. The main reason for the difference in the results is that these high-R&D firms load more negatively on the Fama-French profitability factor, despite earning higher revenues relative to both costs and tangible capital. The risk spanned by these high R&D/Investment firms is not just important for those firms, however. In the Fama-MacBeth test, the long-short portfolio has significant explanatory power for the entire cross-section of excess returns and (along with the market and SMB factors) spans HML such that it is no longer significant. Combined, these results indicate very strongly that high R&D/Investment firms are earning significantly higher returns than those predicted by the leading models of expected returns, and that the risk spanned by these firms is important for the entire cross-section of equity returns.

In seeking to explain this pattern, this paper focused on the exposure of these firms to the high markups which they charge. There is significant evidence that high R&D/Investment firms charge significantly higher markups over cost, even after controlling for industry, size, and year effects. Moreover, the excess returns of these firms correlate significantly with the sales growth of another high-markup item, luxury goods. This effect holds even after controlling for the usual Fama-French factors. Combined, these pieces of evidence suggest that the risk encapsulated by the sales of these high-markup products is important to understand the higher returns of R&D-intensive firms.

The final contribution of this paper is a model which formalizes this intuition. The model integrates a standard production-based asset pricing framework with the novel mechanism of product market interactions between heterogeneous firms and heterogeneous consumers. In the product (final goods) market, heterogeneous firms offer vertically differentiated products to consumers with different levels of willingness to pay for higher quality products. The equilibrium between these two agents results in the purchase of lower-quality goods by consumers who are less willing to pay for quality and the purchase of higher-quality goods at higher markups by consumers who are more willing to pay for quality. The aggregation of consumers choosing a particular firm's product determines the quantity that it decides to produce, but this quantity (and the price that the firm sets) are subject to both supply shocks in the form of productivity shocks and demand shocks in the form of changes to the distribution of consumer preferences. While firms have similar exposures to supply shocks, the model generates the endogenous result that firms offering higher-quality products are more exposed to demand shocks. This risk factor is not spanned by the market risk factor and thus generates excess returns for these firms. The model also matches the size distribution of firms and the markup dynamics of the economy.

While these targets are certainly first-order, one could imagine many other goals for such a model in capturing the dynamics of more sophisticated return models or matching firm leverage or investment timing choices more closely. This product market mechanism seems to be a good first step on the path towards these goals, which are left for future research.

	R&D/Investment				
	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1
Assets ($MM)	4344	2280	2491	1444	448
Leverage (%)	25.3	23.2	20.3	14.8	19.6
PPE/Assets (%)	38.9	28.0	22.0	16.2	10.8
Revenue/Cogs (%)	159.3	169.5	192.0	239.5	246.5
Revenue Growth (%)	21.1	21.1	26.1	33.2	32.9
Revenue/PPE	4.75	6.27	7.73	10.2	16.4
Revenue/Assets	1.19	1.23	1.14	1.02	0.79
Capex/Assets (%)	9.86	7.27	6.02	4.91	2.38
R&D/Assets (%)	0.98	3.15	6.14	12.1	32.9
R&D/Investment (%)	0.10	0.30	0.50	0.70	0.92

Table 4: R&D/Investment Summary Statistics.

Appendix A: Tables and Figures

Note: Table presents mean of each variable by R&D/Investment group. All variables or ratios winsorized at the 1% and 99% levels. Firm-year observation level.

% obs		Time $t+1$ R&D/Inv Decile									
		1	2	3	4	5	6	7	8	9	10
Time t R&D/Inv Decile	1	98%	0.5%	0.5%	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%
	2	3.4%	65%	25%	3.1%	1.2%	0.6%	0.1%	0.5%	0.2%	0.2%
	3	1.3%	6.7%	67%	17%	4.1%	1.7%	0.9%	0.5%	0.3%	0.3%
	4	0.6%	0.7%	17%	51%	19%	6.1%	2.3%	1.5%	0.9%	0.4%
	5	0.4%	0.3%	3.7%	21%	43%	18%	7.3%	3.3%	1.7%	0.8%
	6	0.3%	0.1%	1.4%	5.7%	20%	38%	20%	8.3%	4.1%	1.8%
	7	0.3%	0.1%	0.9%	2.3%	6.9%	21%	34%	21%	9.9%	3.8%
	8	0.1%	0.1%	0.5%	1.2%	3.1%	8.7%	22%	34%	22%	8.2%
	9	0.3%	0.1%	0.4%	0.8%	1.7%	4.1%	9.3%	23%	39%	22%
	10	0.4%	0.0%	0.3%	0.5%	0.9%	1.7%	3.6%	8.0%	22%	62%

Table 5: R&D/Investment Transition Matrix

Note: Annual transition matrix for R&D/Investment deciles. Value in ij^{th} entry represents the probability that a firm in the i^{th} R&D/Investment decile in year t is in the j^{th} decile in year $t+1$. Firm-year observation level.This measure is significantly more persistent than the book-to-market, profitability, investment, and momentum measures. Numbers in each row sum to 100% (with possible rounding error).

Equal-weighted:

Alphas	1	2	3	4	5	6	7	8	9	10	10-1
					R&D/Sales						
α	-0.094	-0.039	0.075	0.113	0.310	0.297	0.394	0.428	0.282	0.269	0.364*
	(-1.05)	(-0.48)	(0.90)	(1.35)	(3.17)	(2.61)	(3.12)	(3.10)	(1.77)	(1.36)	(1.81)
					R&D Capital/Market Equity						
α	-0.436	-0.189	-0.142	-0.013	-0.069	0.170	0.217	0.356	0.712	1.381	1.817***
	(-5.08)	(-2.78)	(-1.93)	(-0.18)	(-0.82)	(1.79)	(2.00)	(2.67)	(4.25)	(6.15)	(7.89)
					R&D/R&D Capital						
α	0.455	0.431	0.350	0.251	0.331	0.177	0.108	0.148	-0.217	-0.102	-0.691***
	(2.94)	(3.46)	(3.95)	(2.96)	(4.12)	(1.97)	(1.11)	(1.10)	(-1.73)	(-0.55)	(-4.08)

Table 6: Equal-Weighted vs. Value-Weighted Results for Other R&D Measures

Value-weighted:

Alphas	1	2	3	4	5	6	7	8	9	10	10-1
					R&D/Sales						
α	-0.010	0.040	-0.040	0.025	0.086	0.084	-0.003	0.413	0.416	0.055	0.154
	(-0.92)	(0.46)	(-0.42)	(0.28)	(0.98)	(0.92)	(-0.03)	(3.12)	(2.53)	(0.27)	(0.62)
					R&D Capital/Market Equity						
α	-0.013	0.003	0.160	0.198	0.134	0.095	0.194	0.145	0.087	0.092	0.105
	(-0.12)	(0.03)	(1.94)	(2.23)	(1.46)	(0.95)	(1.72)	(1.14)	(0.63)	(0.63)	(0.55)
					R&D/R&D Capital						
α	-0.003	0.178	0.092	0.121	0.023	0.108	0.168	0.052	0.084	-0.001	-0.144
	(-0.03)	(1.92)	(1.21)	(1.60)	(0.29)	(1.06)	(1.60)	(0.38)	(0.57)	(-0.01)	(-0.61)

Table 7: Equal-Weighted vs. Value-Weighted Results for Other R&D Measures

Note: Tables present Fama-French 3-factor equal- and value-weighted alphas sorted by deciles of common R&D measures in the literature. Alphas are reported as basis points (bps) per month. In the last column, *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. These results use the same data sample and portfolio construction methodology as the R&D/Investment results presented in the paper. Numbers in parentheses are Newey-West t-statistics.

	R&D/Investment Decile							
	0	1	2	4	6	8	10	10-1
α	-0.130	-0.258	-0.132	0.049	0.266**	0.429***	0.767***	1.025***
	(-1.17)	(-2.84)	(-1.59)	(0.60)	(2.60)	(3.19)	(3.85)	(5.27)
$RMRF$	0.993	1.090	1.059	1.062	1.066	1.045	1.049	-0.041
HML	0.408	0.333	0.323	0.087	-0.225	-0.333	-0.297	-0.630
SMB	0.895	0.569	0.651	0.746	0.948	1.158	1.557	0.988

Table 8: Equal-Weighted R&D/Investment Results

Note: Table reports Fama-French 3-factor equal-weighted alphas and betas by deciles of R&D/Investment measure. Alphas are reported as basis points (bps) per month. In the top row, *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. These results use the same data sample and portfolio construction methodology as the value-weighted results presented in the paper. Numbers in parentheses are Newey-West t-statistics.

	λ Values	
$High - Low$ $\frac{RD}{Inv}$	0.599	0.658*
	(1.32)	(1.87)
$Mkt - rf$	0.594***	0.551**
	(2.91)	(2.34)
HML	0.126	0.249
	(0.79)	(1.39)
SMB	0.063	0.242
	(0.39)	(1.29)
RMW	-0.076	-0.163
	(-0.60)	(-1.20)
CMA	0.174	0.223
	(1.50)	(1.66)
Weights	Value	Equal

Table 9: Industry Fama-Macbeth Results for Fama-French 5 Factors and Innovation Factor

Note: This table reports the Fama-Macbeth prices of risk from the two-stage Fama-Macbeth regression presented in the paper for value- and equal-weighted industry returns. Values reported as percentage points per month. First, for each industry at each point in time, 72-month rolling-window betas with respect to the innovation factor and simultaneously the Fama-French 5-factor model. After that, at each point in time, a cross-sectional regression of excess returns on betas is computed, and the prices of risk extracted as the λ values in specification (4). Specifically, each estimate reported is an estimate of a lambda value based on a time-series average of the lambdas estimated for each cross-sectional second-stage regression. Results robust to value vs. equal weighting of factors and different horizons for rolling window estimation. *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. Numbers in parentheses are t-statistics corrected for autocorrelation over time.

	λ Values	
$High - Low$ $\frac{RD}{Inv}$	0.474*	0.520**
	(1.84)	(2.05)
$Mkt - rf$	0.405**	0.397**
	(2.12)	(2.08)
HML	0.035	0.047
	(0.26)	(0.37)
SMB	0.336**	0.327**
	(2.43)	(2.42)
RMW		-0.193**
		(-2.06)
CMA		0.102
		(1.15)

Table 10: Individual Stock Fama-Macbeth Results

Note: This table reports the Fama-Macbeth prices of risk from the two-stage Fama-Macbeth regression presented in the paper for value-weighted stock returns. Values reported as percentage points per month. First, for each stock at each point in time, 60-month rolling-window betas with respect to the innovation factor and simultaneously the Fama-French 3-factor model or the Fama-French 5-factor model. After that, at each point in time, a cross-sectional regression of excess returns on betas is computed (again, separate regressions for the 3 factors and R&D from the 5 factors and R&D), and the prices of risk extracted as the λ values in specifications (3) and (4). Specifically, each estimate reported is an estimate of a lambda value based on a time-series average of the lambdas estimated for each cross-sectional second-stage regression. *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. Numbers in parentheses are t-statistics corrected for autocorrelation over time.

	$Markup_{t+1}$
$\frac{RD_t}{Inv_t}$	0.424***
	(3.20)
$ln(Assets_t)$	0.048
Industry FE	Yes
Year FE	Yes
Obs.	129,097

Table 11: R&D/Investment and Firm Markup

Note: Markup is defined as Revenue/Cost of Goods Sold -1. *** indicates significant at the 1% level. Numbers reported in parantheses are standard errors clustered at industry level. All variables or ratios winsorized at the 1% and 99% levels. Firm-year observation level.

	$ret_{1,t}$	$ret_{10,t}$	$ret_{10-1,t}$
L_t	-0.278	0.664*	0.942**
	(-1.48)	(1.82)	(2.41)
$RMRF_t$	0.860	1.194	0.334
HML_t	0.209	-0.714	-0.923
SMB_t	-0.102	1.247	1.348
Intercept	-0.114	0.491	0.605
Obs.	192	192	192
$Overall - R^2$	75.32	79.64	69.23

Table 12: R&D/Investment Returns and Luxury Good Sales

Note: L_t measure is real growth in luxury sales from Ait-Sahalia et al. (2004). ret_{10} represents the value-weighted return on the firms in the highest R&D/Investment decile. The coefficient can be interpreted as follows: a 1 standard deviation increase in luxury good sales is associated with a 66 bps increase in the monthly returns of the high R&D/Investment portfolio and a 28 bps decrease in the monthly returns of the low R&D/Investment portfolio, after controlling for the exposure of this portfolio to the Fama-French 3 factors. For L_t, * indicates significant at the 10% level. Numbers reported in parantheses are Newey-West standard errors. Firm-year observation level.

Parameter	Value	Moment/Target
β	-0.006	r_f mean
γ_{a0}	35	r_f std dev
γ_{a1}	-700	r_M Sharpe
γ_{x0}	.5	r_{HI} mean
γ_{x1}	-10	r_{LO} mean
r_k	0.01	physical cap dep
r_{ik}	0.02	intangible cap dep
ρ_a	0.998	
μ_a	0	Zhang (2005)
σ_a	0.002	
ρ_x	0.998	Markup series ρ
μ_x	1.0	Markup series μ
σ_x	0.06	Markup series σ

Table 13: Calibrated Parameters

Note: Calibration is at a monthly frequency. See discussion in Section 4.1.

Moment	Model	Data	Source
r_f mean	0.022	0.018	Campbell (2001)
r_f std dev	0.029	0.030	Campbell (2001)
r_M Sharpe	0.41	0.43	Campbell (2001)
Markup series ρ	0.99	0.9	Corhay, Kung, Schmid (2015)
Markup series μ	0.1413	0.1339	Corhay, Kung, Schmid (2015)
Markup series σ	0.0306	0.0230	Corhay, Kung, Schmid (2015)

Table 14: Model vs. Data Moments

Note: See discussion in Section 4.1. Return moments are presented on a monthly basis as in Zhang (2005) while markup moments are presented at a quarterly frequency as in Corhay, Kung, and Schmid (2015).

	CAPM Results			
	α		β	
	Model	Data	Model	Data
Low R&D/Inv	-0.007	-0.016	1.001	0.925
Medium R&D/Inv	0.11	0.084	0.993	1.074
High R&D/Inv	0.18	0.271	1.065	1.270
	Size			
	Model	Data		
Low R&D/Inv	9.04	9.24		
Medium R&D/Inv	3.26	5.68		
High R&D/Inv	1	1	(normalized)	

Table 15: Model Predictions for Returns and Size

Note: Firms divided into three categories by level of R&D/Investment in both data and model. Size measured as Net PP&E. Alphas are reported as basis points (bps) per month. Model moments taken as mean of 1,000 samples of 600 observations, data counterparts also use 600 observations.

FF3-Alpha		R&D/Investment Quintile					
		1	2	3	4	5	5-1
Assets Quintile	1	-0.652	-0.337	-0.125	-0.259	0.126	0.778***
		(-3.73)	(-1.70)	(-0.56)	(-1.02)	(0.44)	(2.57)
	2	-0.417	0.175	0.341	0.028	0.133	0.550**
		(-2.65)	(1.17)	(1.63)	(0.20)	(0.69)	(2.26)
	3	-0.225	-0.089	0.565	0.423	0.433	0.658***
		(-1.59)	(-0.71)	(3.67)	(2.76)	(2.68)	(3.22)
	4	-0.212	0.091	0.019	0.119	0.558	0.770***
		(-1.96)	(0.75)	(0.17)	(1.00)	(4.35)	(4.60)
	5	-0.047	-0.057	0.045	0.017	0.438	0.484***
		(-0.45)	(-0.64)	(0.51)	(0.19)	(4.94)	(3.08)

Table 16: Doublesorts on Firm Size and R&D Measures

Note: Firms sorted first into quintiles on Assets and then on quintiles based on R&D measures. Table reports Fama-French 3-factor value-weighted alphas by group. Alphas are reported as basis points (bps) per month. In the last column, *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. These results use the same data sample and portfolio construction methodology as the value-weighted results presented in the paper. Numbers in parentheses are Newey-West t-statistics. Insignificance of factor for higher asset quintiles is similar for most other existing R&D measures, including R&D Capital/Market Equity and R&D/R&D Capital.

FF3-Alpha		R&D/Market Equity Quintile					
		1	2	3	4	5	5-1
Assets Quintile	1	-0.909	-0.140	-0.079	0.350	0.894	1.782***
		(-4.35)	(-0.62)	(-0.40)	(1.53)	(2.91)	(4.99)
	2	0.039	0.255	-0.185	0.236	0.602	0.563*
		(0.22)	(1.76)	(-1.29)	(1.44)	(2.66)	(1.96)
	3	0.265	0.238	0.364	0.299	0.428	0.164
		(1.70)	(1.80)	(2.66)	(2.17)	(2.56)	(0.73)
	4	0.229	0.158	0.091	0.172	0.245	0.016
		(1.78)	(1.52)	(0.85)	(1.43)	(1.30)	(0.07)
	5	0.032	0.236	0.058	0.052	0.029	-0.004
		(0.37)	(3.23)	(0.68)	(0.50)	(0.23)	(-0.02)

Table 17: Doublesorts on Firm Size and R&D Measures

Note: Value in ij^{th} entry represents the probability that a firm in the i^{th} R&D/Investment decile in year t is in the j^{th} R&D/Market Equity decile in year t. Firm-year observation level. Numbers in each row sum to 100% (with possible rounding error).

% of R&D/Inv obs		R&D/Market Equity Decile									
		1	2	3	4	5	6	7	8	9	10
R&D/Inv Decile	1	51.06%	19.51%	9.28%	5.14%	3.13%	2.31%	2.01%	1.80%	1.99%	3.78%
	2	20.92%	24.30%	17.67%	12.27%	7.90%	5.25%	3.44%	2.43%	1.76%	4.06%
	3	9.10%	17.45%	18.17%	15.51%	12.25%	9.44%	6.43%	4.50%	3.04%	4.11%
	4	5.21%	11.26%	14.83%	15.07%	13.89%	12.00%	9.49%	6.90%	4.89%	6.46%
	5	3.70%	8.11%	11.14%	13.06%	14.03%	13.44%	11.87%	10.28%	7.71%	6.66%
	6	3.04%	5.96%	8.78%	10.90%	12.74%	13.58%	13.94%	12.49%	10.63%	7.96%
	7	2.14%	4.42%	6.90%	9.13%	11.33%	12.97%	14.17%	14.86%	13.81%	10.26%
	8	1.77%	3.92%	5.50%	7.86%	9.56%	11.55%	13.79%	15.39%	17.54%	13.11%
	9	1.46%	2.84%	4.28%	5.84%	7.95%	10.44%	13.52%	16.44%	19.34%	17.90%
	10	1.89%	2.72%	4.07%	5.74%	7.49%	9.24%	11.38%	14.69%	18.93%	23.86%

Table 18: R&D/Investment and R&D/Market Equity Similarity Matrix

% of R&D/Inv obs		R&D/Sales Decile									
		1	2	3	4	5	6	7	8	9	10
R&D/Inv Decile	1	64.61%	20.71%	8.06%	3.14%	1.60%	0.62%	0.35%	0.23%	0.33%	0.34%
	2	22.51%	35.18%	19.68%	11.26%	5.21%	2.73%	1.35%	0.85%	0.66%	0.58%
	3	6.67%	23.00%	27.07%	19.43%	9.73%	6.04%	3.24%	2.13%	1.57%	1.11%
	4	2.42%	10.17%	20.65%	23.00%	17.91%	11.12%	6.41%	3.92%	2.73%	1.67%
	5	1.34%	4.59%	10.655	17.78%	20.22%	17.15%	12.06%	8.07%	5.08%	3.07%
	6	0.48%	2.27%	5.39%	10.84%	17.45%	19.24%	17.66%	13.64%	8.37%	4.65%
	7	0.30%	1.24%	3.16%	6.06%	12.10%	16.86%	19.65%	19.50%	14.45%	6.68%
	8	0.10%	0.54%	1.82%	3.50%	7.35%	13.18%	19.28%	22.11%	20.77%	11.25%
	9	0.10%	0.40%	1.05%	2.54%	4.78%	8.65%	13.54%	19.39%	25.64%	23.91%
	10	0.25%	0.34%	0.84%	1.62%	2.52%	3.79%	6.31%	10.67%	22.27%	51.37%

Table 19: R&D/Investment and R&D/Sales Similarity Matrix

Note: Value in ij^{th} entry represents the probability that a firm in the i^{th} R&D/Investment decile in year t is in the j^{th} R&D/Sales decile in year t. Firm-year observation level. Numbers in each row sum to 100% (with possible rounding error).

% of R&D/Inv obs		R&D Capital/Market Equity Decile									
		1	2	3	4	5	6	7	8	9	10
R&D/Inv Decile	1	45.11%	19.88%	10.98%	6.60%	3.96%	2.73%	2.43%	2.33%	2.32%	3.65%
	2	19.42%	21.59%	17.23%	13.44%	9.21%	6.14%	4.19%	2.72%	1.89%	4.16%
	3	9.47%	16.09%	16.63%	14.63%	12.79%	10.45%	7.36%	5.13%	3.35%	4.10%
	4	6.58%	10.43%	13.43%	13.98%	13.55%	12.07%	10.18%	7.73%	5.45%	6.60%
	5	4.96%	8.19%	10.76%	12.52%	12.94%	13.04%	12.11%	10.22%	8.41%	6.86%
	6	4.33%	6.77%	8.82%	10.34%	11.83%	12.66%	13.54%	12.67%	10.76%	8.29%
	7	3.14%	5.54%	7.28%	9.09%	10.80%	12.21%	13.36%	14.17%	13.94%	10.45%
	8	2.84%	4.56%	6.00%	7.78%	9.35%	11.34%	13.35%	15.26%	16.72%	12.79%
	9	2.20%	3.84%	4.88%	6.41%	8.30%	10.38%	12.44%	15.32%	18.34%	17.88%
	10	2.39%	3.63%	4.61%	5.72%	7.58%	9.18%	10.99%	14.05%	18.41%	23.43%

Table 20: R&D/Investment and R&D Capital/Market Equity Similarity Matrix

Note: Value in ij^{th} entry represents the probability that a firm in the i^{th} R&D/Investment decile in year t is in the j^{th} R&D Capital/Market Equity decile in year t. R&D Capital calculated as $\overset{[}{\tau}= t - 5]t\sum\left(1 - .2 * (t - \tau)\right) RD_{\tau}$. Firm-year observation level. Numbers in each row sum to 100% (with possible rounding error).

% of R&D/Inv obs		R&D/R&D Capital Decile									
		1	2	3	4	5	6	7	8	9	10
R&D/Inv Decile	1	18.17%	11.07%	10.29%	8.42%	7.57%	7.50%	8.35%	8.18%	12.01%	8.43%
	2	10.99%	10.27%	12.17%	10.68%	10.82%	9.79%	9.61%	8.25%	11.20%	6.22%
	3	8.57%	9.77%	11.43%	12.27%	11.90%	11.21%	10.05%	8.71%	10.75%	5.34%
	4	7.85%	9.01%	10.50%	12.46%	12.24%	11.55%	11.20%	8.96%	11.29%	4.93%
	5	7.31%	9.29%	9.69%	11.11%	11.72%	10.76%	11.62%	10.64%	12.46%	5.39%
	6	7.27%	8.87%	9.95%	9.84%	11.30%	10.96%	12.03%	11.48%	12.82%	5.48%
	7	7.57%	9.29%	9.18%	9.56%	10.72%	10.98%	11.83%	11.38%	13.74%	5.75%
	8	7.52%	9.69%	9.44%	9.57%	10.50%	9.97%	11.59%	11.68%	14.62%	5.42%
	9	10.05%	10.81%	9.27%	8.41%	8.84%	9.28%	10.49%	11.28%	14.94%	6.63%
	10	15.02%	11.60%	8.55%	7.19%	6.69%	7.84%	9.12%	9.94%	15.73%	8.22%

Table 21: R&D/Investment and R&D/R&D Capital Similarity Matrix

Note: Value in ij^{th} entry represents the probability that a firm in the i^{th} R&D/Investment decile in year t is in the j^{th} R&D/R&D Capital decile in year t. R&D Capital calculated as $\overset{[}{\tau} = t-5]t\sum(1-.2*(t-\tau))\,RD_{\tau}$. Firm-year observation level. Numbers in each row sum to 100% (with possible rounding error).

% of R&D/Inv obs		SG&A Capital/Assets Decile									
		1	2	3	4	5	6	7	8	9	10
R&D/Inv Decile	1	43.24%	15.88%	9.86%	7.01%	6.00%	4.87%	4.24%	3.13%	3.06%	2.73%
	2	20.99%	18.94%	14.06%	11.45%	8.79%	7.13%	5.70%	4.78%	4.13%	4.02%
	3	11.28%	16.66%	14.81%	14.17%	11.40%	9.35%	7.08%	5.99%	5.71%	3.55%
	4	7.46%	13.59%	15.11%	13.85%	13.38%	10.42%	8.74%	7.25%	5.88%	4.33%
	5	5.15%	11.23%	12.17%	13.76%	12.58%	12.21%	10.01%	9.20%	8.30%	5.38%
	6	3.07%	7.47%	10.73%	11.90%	12.85%	12.60%	13.52%	11.71%	9.67%	6.49%
	7	2.63%	5.53%	8.08%	9.49%	11.21%	13.03%	14.41%	14.44%	12.06%	9.12%
	8	1.80%	3.96%	6.50%	8.33%	9.89%	12.28%	13.65%	15.54%	15.49%	12.56%
	9	1.85%	3.18%	4.46%	6.07%	8.07%	10.61%	12.48%	15.83%	18.18%	19.26%
	10	2.22%	3.14%	3.93%	4.09%	6.11%	7.39%	9.92%	12.29%	17.97%	32.94%

Table 22: R&D/Investment and Organizational Capital Equity Similarity Matrix

Note: Value in ij^{th} entry represents the probability that a firm in the i^{th} R&D/Investment decile in year t is in the j^{th} SG&A Capital/Assets decile in year t. Organizational (SG&A) Capital calculated as $\overset{[}{\tau}= t - 5]t\sum (1 - .2 * (t - \tau))\, SGA_{\tau}$. Firm-year observation level. Numbers in each row sum to 100% (with possible rounding error).

FF3-Alpha		R&D/Investment Quintile				
		1	2	3	4	5
Gross Profitability Quintile	1	-0.696	-0.843	-0.802	0.164	-0.695
		(-2.43)	(-3.14)	(-2.85)	(0.52)	(-2.24)
	2	-0.035	0.022	0.294	0.422	0.326
		(-0.31)	(0.15)	(1.75)	(2.28)	(1.49)
	3	-0.125	-0.087	0.197	0.501	0.581
		(-1.03)	(-0.85)	(2.09)	(4.02)	(3.21)
	4	0.081	-0.073	0.134	0.425	0.437
		(0.67)	(-0.63)	(1.20)	(3.20)	(2.22)
	5	0.259	0.365	0.194	0.067	0.529
		(1.85)	(2.23)	(1.05)	(0.28)	(2.23)

Table 23: FF3 Doublesorts on Gross Profitability and R&D/Investment

Note: Firms sorted first into quintiles on Gross Profitability and then on quintiles based on R&D/Investment. Gross Profitability defined as Revenue/Assets. Table reports Fama-French 3-factor and 5-factor value-weighted alphas by group. Alphas are reported as basis points (bps) per month. These results use the same data sample and portfolio construction methodology as the value-weighted results presented in the paper. Numbers in parentheses are Newey-West t-statistics.

FF5-Alpha		R&D/Investment Quintile				
		1	2	3	4	5
Gross Profitability Quintile	1	0.010	-0.306	-0.419	0.437	-0.512
		(0.04)	(-1.12)	(-1.52)	(1.42)	(-1.65)
	2	0.040	0.426	0.547	0.633	0.612
		(0.35)	(3.07)	(3.09)	(3.19)	(2.87)
	3	-0.162	-0.051	0.215	0.645	0.839
		(-1.28)	(-0.48)	(2.14)	(5.00)	(4.73)
	4	-0.077	-0.134	0.077	0.437	0.698
		(-0.64)	(-1.11)	(0.63)	(3.24)	(3.46)
	5	0.121	0.230	0.267	0.107	0.721
		(0.87)	(1.31)	(1.30)	(0.44)	(2.93)

Table 24: FF5 Doublesorts on Gross Profitability and R&D/Investment

Note: Firms sorted first into quintiles on Net Profitability and then on quintiles based on R&D/Investment. Net Profitability defined as Net Income/Assets. Table reports Fama-French 3-factor and 5-factor value-weighted alphas by group. Alphas are reported as basis points (bps) per month. These results use the same data sample and portfolio construction methodology as the value-weighted results presented in the paper. Numbers in parentheses are Newey-West t-statistics.

FF3-Alpha		R&D/Investment Quintile				
		1	2	3	4	5
Net Profitability Quintile	1	-0.405 (-1.86)	-0.171 (-0.72)	0.188 (0.80)	-0.525 (-2.22)	0.039 (0.15)
	2	-0.469 (-3.15)	-0.180 (-1.12)	-0.061 (-0.32)	0.033 (0.17)	0.365 (1.65)
	3	-0.331 (-2.57)	-0.010 (-0.08)	-0.097 (-0.74)	0.107 (0.67)	0.232 (1.23)
	4	0.096 (0.76)	-0.212 (-2.30)	0.104 (0.86)	0.137 (1.24)	0.359 (2.44)
	5	0.014 (0.12)	0.092 (0.85)	0.351 (3.27)	0.336 (3.32)	0.588 (3.88)

Table 25: FF3 Doublesorts on Net Profitability and R&D/Investment

FF5-Alpha		R&D/Investment Quintile				
		1	2	3	4	5
Net Profitability Quintile	1	0.038 (0.17)	0.431 (1.96)	0.624 (2.77)	-0.185 (-0.74)	0.302 (1.19)
	2	-0.368 (-2.38)	-0.007 (-0.05)	0.170 (0.89)	0.346 (1.79)	0.707 (3.16)
	3	-0.280 (-2.12)	-0.074 (-0.54)	-0.080 (-0.58)	0.301 (1.85)	0.411 (2.03)
	4	0.094 (0.71)	-0.291 (-3.11)	0.071 (0.58)	0.155 (1.37)	0.496 (3.22)
	5	-0.145 (-1.19)	0.004 (0.03)	0.385 (3.43)	0.287 (2.72)	0.719 (4.61)

Table 26: FF5 Doublesorts on Net Profitability and R&D/Investment

FF3-Alpha		R&D/Investment Quintile				
		1	2	3	4	5
Scaled Investment Quintile	1	-0.559	0.259	-0.187	-1.102	-0.712
		(-2.14)	(1.03)	(-0.77)	(-3.57)	(-1.71)
	2	-0.435	0.161	0.185	-0.270	0.091
		(-2.83)	(0.78)	(1.22)	(-1.52)	(0.39)
	3	-0.070	0.038	-0.025	0.121	0.343
		(-0.58)	(0.31)	(-0.21)	(0.86)	(2.00)
	4	0.006	-0.077	0.101	0.477	0.315
		(0.05)	(-0.83)	(1.06)	(3.65)	(1.76)
	5	0.182	0.255	0.319	0.043	0.771
		(1.10)	(2.00)	(2.04)	(0.23)	(3.57)

Table 27: FF3 Doublesorts on Firm Scaled Investment and R&D/Investment

Note: Firms sorted first into quintiles on Scaled Investment and then on quintiles based on R&D/Investment. Scaled Investment defined as Investment/Assets. Table reports Fama-French 3-factor and 5-factor value-weighted alphas by group. Alphas are reported as basis points (bps) per month. These results use the same data sample and portfolio construction methodology as the value-weighted results presented in the paper. Numbers in parentheses are Newey-West t-statistics.

FF5-Alpha		R&D/Investment Quintile				
		1	2	3	4	5
Scaled Investment Quintile	1	-0.478	0.130	-0.465	-1.004	-0.513
		(-1.82)	(0.50)	(-1.97)	(-2.96)	(-1.18)
	2	-0.449	-0.034	-0.009	-0.216	0.363
		(-2.92)	(-0.15)	(-0.06)	(-1.18)	(1.40)
	3	-0.136	-0.039	-0.048	0.255	0.518
		(-1.06)	(-0.29)	(-0.39)	(1.69)	(2.78)
	4	0.082	-0.037	0.116	0.596	0.467
		(0.76)	(-0.39)	(1.11)	(4.42)	(2.57)
	5	0.429	0.381	0.537	0.265	1.040
		(2.70)	(2.89)	(3.49)	(1.36)	(4.90)

Table 28: FF5 Doublesorts on Firm Scaled Investment and R&D/Investment

	R&D/Investment Decile										
	0	1	2	3	4	5	6	8	10	10-1		
α	-0.234	0.081	0.027	0.085	0.011	0.201**	0.142	0.679***	0.973***	0.892***		
	(-1.57)	(0.81)	(0.28)	(0.85)	(0.12)	(2.04)	(1.34)	(2.72)	(4.44)	(2.38)	(4.92)	(3.80)
$RMRF$	1.195	0.835	1.065	1.038	1.022	0.974	0.979	0.954	0.899	0.083		
HML	0.241	0.112	0.154	-0.042	-0.151	-0.331	-0.333	-0.459	-0.669	-0.931	-0.937	-1.049
SMB	0.431	-0.304	-0.158	0.067	-0.017	-0.027	0.066	0.122	0.121	0.413	0.595	0.898
QMJ	0.458	-0.229	0.002	-0.034	-0.062	-0.181	0.130	-0.045	-0.346	-0.528	-0.746	-0.758

Table 29: Fama-French 3-factor and QMJ Results

Note: Table reports value-weighted alphas and betas by deciles of R&D/Investment measure after controlling for Fama-French 3 factors and QMJ factor of Asness, Frazzini, and Pedersen (2014). Alphas are reported as basis points (bps) per month. In the top row, *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. These results use the same data sample and portfolio construction methodology as the value-weighted results presented in the paper. Numbers in parentheses are Newey-West t-statistics.

	R&D/Investment Decile											
	0	1	2	3	4	5	6	7	8	9	10	10-1
α	0.065	-0.076	0.021	0.129	0.052	0.150	0.196**	0.281**	0.512***	0.087	0.444**	0.520**
	(0.47)	(-0.79)	(0.24)	(1.43)	(0.59)	(1.56)	(2.11)	(2.41)	(3.50)	(0.53)	(2.38)	(2.41)
$RMRF$	1.081	0.894	1.065	1.032	1.020	1.005	0.953	1.027	1.027	1.027	1.112	0.219
HML	0.145	0.162	0.156	-0.059	-0.167	-0.317	-0.348	-0.443	-0.617	-0.826	-0.766	-0.928
SMB	0.275	-0.226	-0.158	0.078	0.005	0.035	0.022	0.137	0.238	0.592	0.847	1.073
MOM	-0.024	0.020	0.007	-0.069	-0.084	-0.062	0.028	0.022	-0.045	0.012	-0.085	0.064

Table 30: Fama-French 3-factor and Momentum Results

Note: Table reports value-weighted alphas and betas by deciles of R&D/Investment measure after controlling for Fama-French 3 factors and Momentum (based on 2-12 month prior return). Alphas are reported as basis points (bps) per month. In the top row, *** indicates significant at the 1% level, ** indicates significant at the 5% level, and * indicates significant at the 10% level. These results use the same data sample and portfolio construction methodology as the value-weighted results presented in the paper. Numbers in parentheses are Newey-West t-statistics.

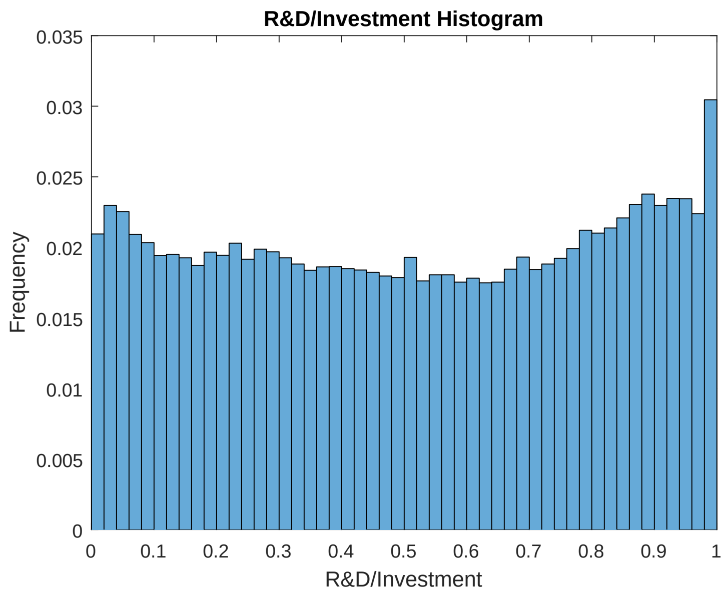

Figure 5: Histogram of R&D/Investment firm-year level observations for reported and nonzero values of R&D

Note: Figure plots the histogram of R&D/Investment firm-year level observations for reported and nonzero values of R&D. Investment is defined as the sum of R&D expenditures and capital expenditures; see Section 2.1 for more details. Approximately 50% of the firm-year observations in the merged sample have missing R&D values, approximately 17% of those with non-missing values have zero values. Firm-year observation level.

Figure 5: R&D/Investment Distribution

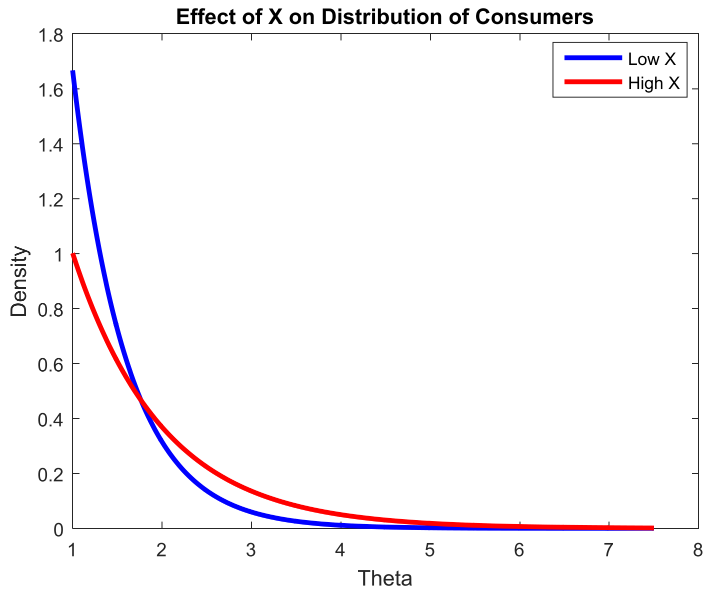

Figure 6: Distribution of consumer preferences in high and low demand states
Note: Figure plots the distribution of consumers' theta prefences under two values of the demand state variable X_t. Lower values of X_t (blue line) correspond to distributions which skew towards lower theta values and have fewer high theta values.

Figure 6: Effect of X on Distribution of Consumer Preferences

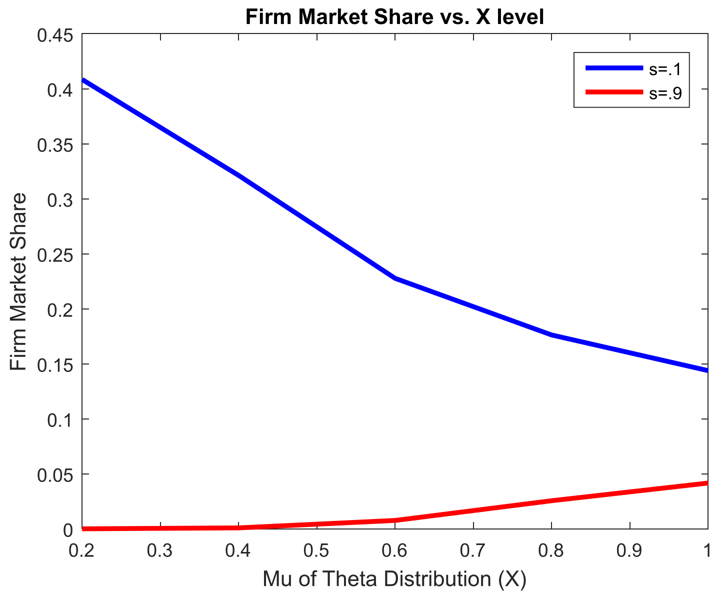

Figure 7: Model-Implied Market Shares vs. Demand State
Note: Figure plots the model-generated market shares of low- and high-quality firms as the preference parameter (X_t) changes. Market shares calculated as the quantity of goods produced and sold by a given firm divided by the quantity produced and sold by all firms. Productivity state fixed to long-run mean.

Figure 7: Model Market Shares

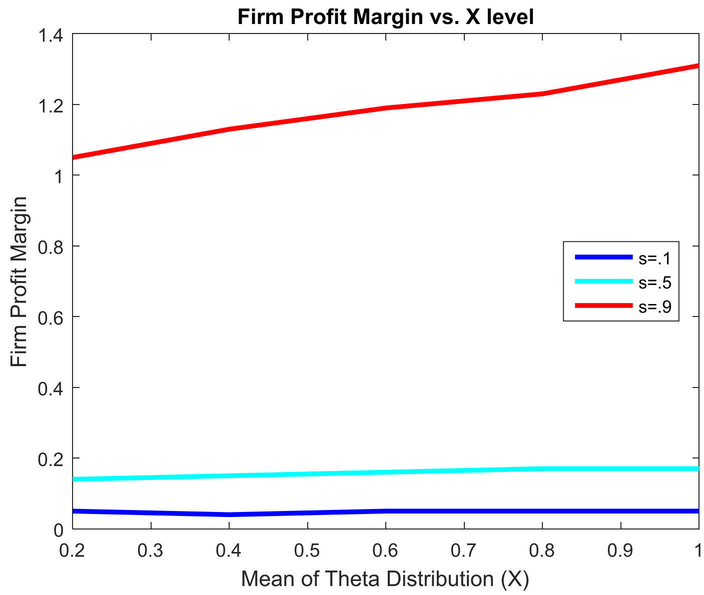

Figure 8: Model-Implied Profit Margins vs. Demand State
Note: Figure plots the profit margins of low- , medium-, and high-quality firms as the preference parameter (X_t) changes. Profit margin calculated as the total profit of a firm (revenue less rental costs of capital) divided by the quantity sold. Productivity state fixed to long-run mean.

Figure 8: Model Profit Margins

Appendix B: Proof of Proposition 1

Proposition 1 can be proved by establishing Lemmas 2 and 3.

Lemma 2. *In any pure-strategy Nash Equilibrium, there will be at most one firm producing products of a given quality level* $s_i \, \forall i$ *.*

Proof. Consider first the problem of a firm seeking to enter that previously did not produce products of any quality level. Since this firm is switching quality levels (in the sense that it previously had no quality level and is now entering a quality level) it must pay some switching costs, denoted by c (without loss of generality, assume that this cost is the same for firms switching from no previous quality level as it is for firms switching from a different previous quality level.) If this firm enters a quality level s_j where there is an existing firm producing, then both this new entrant and the existing firm will decide to produce the same quantity of products, because these firms are identical. Given that these identical firms are competing for the production of a homogeneous good, they will compete on the prices they set such that both firms earn zero profits in equilibrium. Therefore, the payoff to the entrant from entering any quality level which has an existing firm will be $-c < 0$. Since this is less than the payoff for not entering at all (0), this new entrant will never choose to enter a quality level which has an existing firm.

The comparison for a firm which is producing products of a different quality level s_i is even starker. By not switching quality levels, this firm earns profits π_i, which are always weakly positive since the firm can choose to produce zero units of goods and rent zero units of capital and thus earn zero profits. The analysis for the profits of the firm if it enters a quality level where there is already an existing firm is the same as above, and this firm will earn $-c < 0$. Since $-c < 0 \leq \pi_i$, this firm is always strictly better off by not switching to a quality level where there is an existing firm producing goods.

Therefore, no firm will enter or switch into a quality level where it believes that there will be another firm. Since these beliefs are correct in equilibrium, in any pure-strategy Nash Equilibrium, firms will have perfect knowledge of the quality levels occupied by every other firm and will not choose to produce products in any of these occupied quality levels. Thus, in any pure-strategy Nash Equilibrium, there will never be more than one firm producing products of any given quality level.

□

Lemma 3. *Among all possible Nash Equilibria, the initial allocation where firm j produces quality level s_j is Pareto-optimal.*

Proof. In any pure-strategy Nash Equilibrium, there will always be exactly one firm per quality level. This is because each quality level is initially occupied by one firm, and this firm earns (weakly) positive profits in every period. As a result, each of these firms has no incentive to exit, and so will always remain in the economy. Then, as long as there are at least N firms in the economy (the initial number) and N different quality levels, if there is no more than one firm producing products of a given quality then each quality level will have no fewer than one firm, by the Pigeonhole Principle. Combining this with Lemma 2, each quality level will have exactly one firm in a pure-strategy Nash Equilibrium.

As a result, any pure strategy Nash Equilibrium consists of an assignment of N firms to N quality levels such that each quality level has exactly one firm. The initial allocation of firms such that firm j produces quality level s_j is one such equilibrium. This equilibrium can be shown to be Pareto optimal as follows. In this equilibrium, the N firms earn profit levels $\pi_1, \pi_2, \ldots, \pi_i, \pi_j, \ldots, \pi_N$. Any pure strategy modification would result in the switching of two profit levels, and the deduction of c from each. If one profit level is higher than the other, then the firm which moves from the higher to the lower profit level will necessarily be worse off. If the two profit levels are equal, then the switch results in both firms paying switching costs and both being worse off as their profits have been reduced by c. Thus there exist no pure-strategy modifications to the initial allocation of firms that make one firm better without making one firm worse off. Note that this is not necessarily the case with other equilibria because if those equilibria involve the switching of firms, then reducing those switching costs can potentially make multiple firms better off without harming other firms. For example, if $\pi_i = \pi_j$ and the Nash Equilibrium involved firms i and j switching quality levels, then a Pareto improvement on this would be for firm i to stay at quality level s_i and firm j to stay at quality level s_j.

There also exist mixed-strategy Nash Equilibria where some firms have probabilities of occupying various quality levels. Note that none of these equilibria represent a Pareto improvement over the initial allocation either. This can be seen in many ways, but one way is to examine the aggregate profit. The aggregate profit under any mixed-strategy Nash Equilibrium will be weakly lower than that under the initial allocation. This is explained by several components. First, since firms make identical quantity decisions once they are put into a given quality level, if there are quality levels with only one firm under the mixed-strategy Nash Equilibrium, then the firms in those quality levels will earn the same profits as the firms in those quality levels under the initial allocation, less any switching costs. Therefore the profits of these firms will be weakly lower. Second, if there are quality levels with multiple firms, the firms in those quality levels will earn zero profits, again weakly lower than under the initial allocation. If the mixed-strategy Nash Equilibrium involves any switching at all–that is, if there are any firms that use mixed strategies–then the aggregate profits will be strictly lower. In such a mixed-strategy Nash Equilibrium, these lower profits can be split among a greater, lower, or equal number of firms. If the lower profits are split among a greater or equal number of firms, then it is clear that some firms must be worse under this equilibrium than in the initial allocation. If they are split among fewer firms, then there are some firms that would earn positive profits under the initial allocation (since all firms under the initial allocation earn strictly positive profits) who now earn zero profits. □

CHAPTER 2 : **Why Do Firms Issue Callable Bonds?**

2.1. **Introduction**

One of the most salient trends in corporate debt markets in the past 10-15 years has been the increasing prevalence of callable bonds (see figure 1). In this paper we study and explain this increase. We find that this increase is driven primarily by a specific type of callable bond, in which the call option of the firm to repurchase the debt is almost never in the money (that is, the strike price of the option is almost always higher than the value of the bond.) With this in mind, we evaluate whether previous motives related to interest rate risk management, asymmetric information, or agency issues can explain the popularity of these bonds. We find that they cannot. Based on relatively new theories and evidence proposed by Mian and Santos (2011) and Xu (2016) and others, we suggest a new motivation for why these bonds occur. Specifically, we show that the presence of these "out of the money" options can help mitigate refinancing or rollover risk for the firm. We propose a simple model featuring this mechanism and use it to help explain several of the empirical patterns that we observe (and document) over the last decade.

While initially popular in the 1980s for interest rate management purposes, the usage of callable bonds declined significantly with the widespread availability of OTC derivatives in the early 1990s. Academic literature around that time suggested that the primary purpose of call options in debt were to help alleviate agency conflicts or problems of asymmetric information (see e.g. Crabbe and Helwege (1994)). In the late 1990s, the usage of callable bonds began to increase, and soon the majority of bonds issued by nonfinancial corporations contained call provisions. This trend has increased over time, to the point where over 90%

of bonds issued by nonfinancial corporations in our sample contained call provisions in each of the last 5 years. This increase has occurred across almost all types of bonds and firms, although there are some cross-sectional differences, as we will discuss in Section 3.

An equally interesting element of this pattern is that the usage of the sorts of callable bonds that were popular in the 1980s peaked in 1999 and fell dramatically thereafter, representing less than 10% of the total par value of bonds issued in 2012-2014. Instead, this increase was driven by a different type of callable bond: the make-whole bond. While substantially similar in terms and structure to the callable bond, the make-whole bond contains one very important difference: the way that the strike price of the option is computed. In particular, the strike price is computed in such a way so as to almost never be below the market value of the bond. Thus, firms exercising the option on these bonds would almost always be doing so at strike prices that are "out of the money." Figure 2 plots the trends in both make-whole bonds and non-make-whole callable bonds for the firms in our sample.

This important distinction has several implications. First, as we will show in Section 4, these make-whole provisions completely remove any interest rate motives for having an embedded call option. Moreover, the agency and asymmetric information stories that potentially explain why callable bonds were issued in the 1990s cannot explain the issuance of make-whole bonds. The structure of the strike prices precludes the manager/equityholders from having the proper incentives and several of their empirical predictions do not hold for the new class of make-whole bonds.

Since these theories do not hold for the new bond structures and are not well-supported by the new empirical evidence, we begin our proposal of a new theory of the issuance of make-whole debt by documenting a few novel stylized facts. First, we show that firms that are likely to have higher credit risk (as measured by income volatility or leverage) are more likely to issue make-whole bonds. Second, we show that, compared to the proceeds of other bonds, firms are more likely to use the proceeds of make-whole bonds for the purposes of investment. Based on these two facts, we propose a theory for the issuance of make-whole bonds where the primary reason that firms prefer bonds of this type is to avoid the potential for refinancing risk. This risk is a topic that has been explored in several other recent papers (e.g. Xu (2016)) and is highest for firms with high credit risk and who invest immediately and so face a potential mismatch between investment and debt maturity. We embed this refinancing risk in a simple model and show that make-whole bonds can help firms previously frozen out of credit markets access capital and allow other firms to borrow more.

This paper is related to several strands of literature. First, it draws upon the early reasons advanced for the usage of callable bonds by firms. One such reason is interest rate risk management by firms, which is discussed in Kraus (1973), building upon the work of Kalyman (1971) and Weingartner (1967), among others. Given that the motivation for interest rate risk management through bond options became somewhat moot with the introduction of OTC interest rate derivatives, new explanations were needed. Many of these centered around agency conflicts; Crabbe and Helwege (1994) gives an excellent overview of this. They identify three primary theories. The first is the problem of managers under-investing if equityholders do not benefit from the returns, a problem identified by Myers (1977). Bodie and Taggart (1978) show that callable debt can help resolve this problem. Barnea, Haugen, and Senbet (1980) identify two further potential agency conflicts: the first stemming from asymmetric information (as also discussed in Myers and Majluf (1984)) and the second from risk-shiftng on the part of managers. These theories will be discussed in more detail in Section 4. The empirical work testing these theories is also quite relevant as it helps identify testable predictions and testing methodologies. Thatcher (1985), Mitchell (1991) and Kish and Livingston (1992) were among the early works to test these hypotheses. More recently, Banko and Zhou (2010) and Guntay, Prabhala, and Unal (2013) have tested some of these hypotheses. Our paper will differ from both groups in that we use the additional characteristic of the more recent bonds as make-whole to more rigorously test these hypotheses (and also in that we use a more comprehensive and newer set of data.)

After a thorough analysis of existing explanations for callable debt, we then propose our own explanation for the prevalence of make-whole bonds relying on the ideas of refinancing risk and maturity management. Although the idea of matching the maturities of assets and debt is a fairly well-established one (see e.g. Modigliani and Sutch (1966) and Myers (1977)), the idea of rollover risk impacting firm decisions is one that is only now gaining much attention. Recent paper such as Brunnermeier and Yogo (2009), Mian and Santos (2011), and Xu (2016) study this extensively. The last of these three, by discussing the impact of callable debt, is particularly related to this paper. This model also builds upon the more canonical models of debt dynamics, such as Leland and Toft (1996), although the structure is a bit different.

The next section will discuss precisely what the differences between make-whole and traditional callable debt are and present evidence on other forms of early refinancing. Section 3 then discusses the data used in the paper and presents several empirical trends, notably on the cross-sectional differences in debt issuance and on the use of proceeds from debt issuance, that will be useful in motivating the model. Following that, Section 4 builds upon Sections 2 and 3 by showing how previous explanations for the usage of callable debt run afoul of either the new institutional characteristics of make-whole debt or the more recent empirical trends. Section 5 presents an alternative model for make-whole debt relying upon incomplete capital markets, refinancing risk, and costs of financial distress. The results of this model are also presented in this section. Finally, Section 6 concludes.

2.2. **Institutional Background**

This section covers two important pieces of institutional background. The first subsection discusses the difference between traditional callable bonds and make-whole bonds and the second subsection reviews the other methods by which a firm may retire its debt early.

2.2.1. Callable and Make-Whole Bonds

The feature differentiating callable bonds from noncallable bonds is a call provision, which the issuer of the debt (in this case the firm) can exercise to repurchase its debt from bondholders. This call provision contains a number of important details. First, it specifies a window during which the bond may be called. This window may be from issuance until the maturity of the bond or only cover a subset of the time that the bond is outstanding. Second, like a traditional call option, the call provision specifies a "strike price" at which the bond may be called. This is where the key difference between traditional callable and make-whole bonds comes, and so we examine it in further detail. [1]

For purposes of our analysis, we consider two primary classes. The first is the case of traditional callable bonds, such as those that were issued throughout the 1980s. Nearly all callable debt issued until 1994 was of this form. These bonds specify a call price (expressed as a percentage of par) at which a firm may call the bond. This price is typically either fixed or varies with time (usually decreasing monotonically) over the length of the call window. Importantly, this is the only dimension along which the price can vary. That is, the path of the strike price of the call option requires depends only on the date. An example of a bond with this traditional call provision comes from Wells Fargo's 17-year \$13.7 million notes issued on June 17, 2014, which state:

"The notes are redeemable by Wells Fargo, in whole or in part, on any interest payment date occurring on or after June 17, 2019 at 100% of their principal amount plus accrued and unpaid interest to, but excluding, the redemption date."

The second class consists of make-whole bonds, which have a strike price structure that has one important change. For make-whole calls, the strike price is set to be the maximum of

[1] There may be other features included in call provisions, including multiple tiers of calls which specify different prices for different date ranges of calls (and in some cases also restrict the number of bonds that can be called) and provisions which specify certain conditions under which a bond issuer may or may not call (these may be either firm specific or macroeconomic conditions,) but these are not highly prevalent. There does, however, exist some literature on firms optimally choosing call provisions to reduce agency costs, among other things (see e.g. Thatcher 2005.) We largely abstract away from these considerations.

the par value (or some fixed percentage of the par value) and a proxy for the market value of the bond. This market value proxy is computed by taking the remaining interest and principal payments of the bond and discounting them at a fairly low interest rate, usually given by a benchmark Treasury rate plus some fairly low fixed spread. It is important to note that this fixed spread is usually set to be below whatever spread the firm could borrow at in the open market, even under the best conditions. An example of a bond with this provision is Coca Cola's November 1, 2013 issue of four fixed rate bonds due in 2016, 2018, 2020, and 2023:

> "We may redeem any series of fixed rate notes at our option and at any time, either as a whole or in part. If we elect to redeem a series of fixed rate notes, we will pay a redemption price equal to the greater of:
>
> - 100% of the principal amount of the notes to be redeemed, plus accrued and unpaid interest; and
> - the sum of the present values of the remaining scheduled payments, plus accrued and unpaid interest.
>
> In determining the present value of the remaining scheduled payments, we will discount such payments to the redemption date on a semi-annual basis (assuming a 360-day year consisting of twelve 30-day months) using a discount rate equal to the Treasury rate plus 5 basis points for the 2016 notes, a discount rate equal to the Treasury rate plus 7 basis points for the 2018 notes, a discount rate equal to the Treasury rate plus 10 basis points for the 2020 notes and a discount rate equal to the Treasury rate plus 10 basis points for the 2023 notes."

This has two immediate implications. First, it means that the strike price will vary not only with time but with market conditions. In particular, the strike price will be highest when Treasury rates are low (generally in good times) and will be lowest when Treasury rates are high. Hence, the strike price will be procyclical. This feature is designed to ensure that the calling of a bond does not expose bondholders to losses due to changes in the market interest rates over the lifetime of the bond. Second, because the fixed spread to Treasuries is set to be below the spread at which a firm could realistically refinance its debt, this price will virtually never be below the market value of the bond. Put another way, if the firm were to reissue a bond with the exact same interest payments and principal, it would almost certainly receive less than it would have to pay to call the identical make-whole bond. This has the effect, as alluded to earlier, of making this call option almost never "in the money" in the sense that the strike price for this call option will almost always be above the market value of the underlying asset (in this case the remaining payments of the bond.)

The following figure gives an example of this. Consider a firm that has issued a five year bond at par with annual coupon payments of 5.5% of the principal ($100) and wants to refinance this bond at year two. Since we know that the make-whole fixed spread is usually far lower than the credit spread at which the firm is reissuing the bond, let's assume that the credit spread at which the firm reissues is 150bps and that the fixed spread to the benchmark Treasury that the firm has to pay as stipulated in the make-whole provision is 30bps. The following plot gives the prices that the firm would have to pay to call a traditional bond (we assume that the call price is fixed to par) and a make-whole bond, as well as the proceeds that the firm would earn from reissuing a bond with exactly the same remaining payoffs as the retired bond. All lines are plotted versus the underlying Treasury rate.

Here, the firm profits by calling its debt and refinancing when the proceeds from reissuance (green line) exceed the price paid to call the debt (red line for traditional callable debt, blue line for make-whole debt.) Note also that this is a zero-sum game: any gain the firm makes by calling the debt below its true value is lost by the bondholder who has to surrender an asset for less than it is worth. Here we see that, for sufficiently low interest rates, the firm can profit if it has traditional callable debt by calling its debt at par and then reissuing debt with the same payments for a higher value. Upon closer inspection, one sees that once the Treasury rate drops below the coupon of the bond less the reissuance credit spread, calling

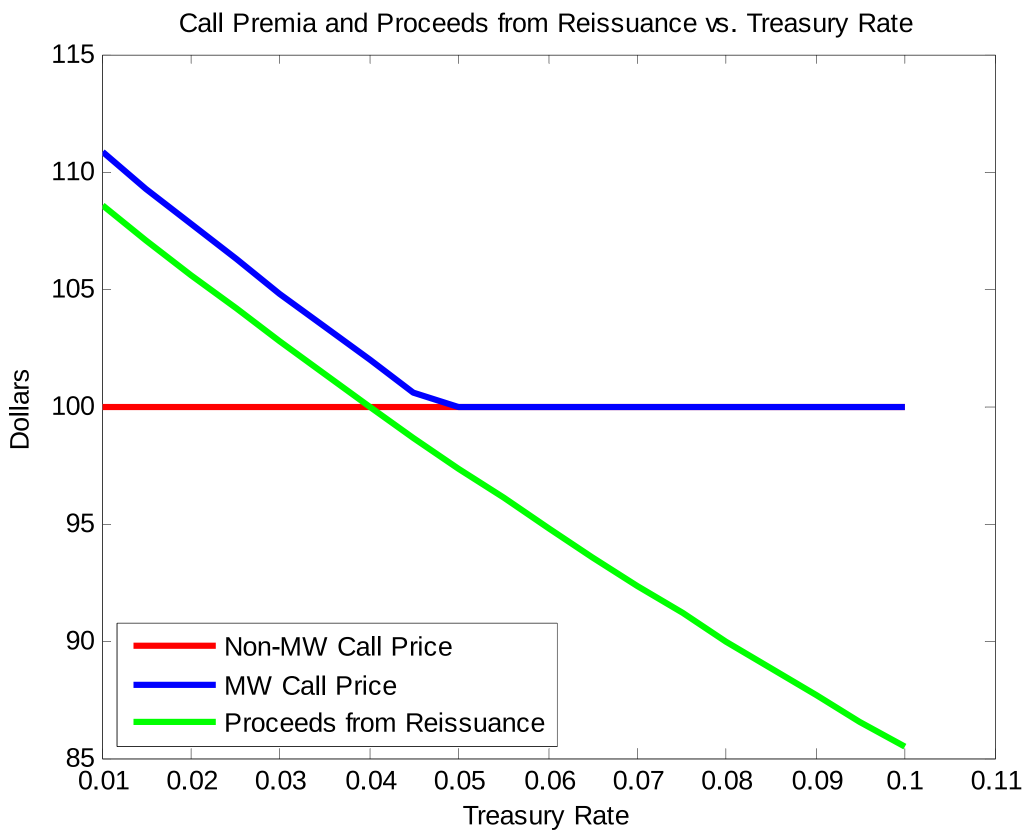

Figure 9: Proceeds from Calling Make-Whole and Traditional Bonds

debt at par becomes valuable for the firm. This is because the interest rate at which that payment is discounted (the Treasury rate +150bps credit spread) is lower than the coupon payment for Treasury rates less than 4%, meaning that the price of the bond is higher than par.

In contrast, it is never profitable for the firm to exercise the call option on make-whole debt and reissue its debt in this example since the credit spread at which it would have to reissue its debt is higher than the spread that it would have to pay to buy back its debt. This means that the discount rate that the firm uses to value the payments that it buys back will be lower than those that will be used to value the debt that it issues and generate the proceeds for the firm. Thus, the make-whole call price will always exceed the proceeds from reissuance. Note also that since the make-whole call price is calculated as the maximum of the traditional call price and the proxy for market value, it will always result in a (weakly) higher payment by the firm.

These make-whole payment spreads are set such that they are almost always below the reissuance credit spread for a firm and thus the firm cannot benefit by calling its make-whole debt and reissuing a bond with the exact same payment structure. Indeed, for most of the bonds in the sample the make-whole spreads range between 5 and 50bps, far lower than the average issuance credit spread for bonds in the sample. Nevertheless, it is worthwhile to consider the extreme case in which the reissuance credit spread is actually lower than the make-whole spread. Figure 3 presents a graph identical to the one above, except that the reissuance credit spread is now 0bps. Note that, since the make-whole spread is so low, the firm still requires extremely low interest rates to be able to make even a slight profit by calling its make-whole debt and reissuing an identical bond.

2.2.2. Other Forms of Early Retirement

In addition to call provisions, there are two other mechanisms by which firms can retire debt early that bear mentioning.

First, a firm can perform an open market repurchase.Transactions such as this are typically executed between two dealers, over the phone, with neither party knowing who the other party represents. For example, a firm may contact an investment bank, asking it to contact pension funds or insurance companies that hold its debt to buy the debt from them. The price is then privately negotiated between the two dealers. There are a couple of important points to note about transactions such as this. First, as noted in Levy and Shalev (2013), "corporate bond transactions [of this type] are relatively sparse." In our complete data sample, open market repurchases constitute 3.57% of all early bond retirements. Second, given that this is done on a bondholder-by-bondholder basis, it is difficult to retire a significant fraction of the outstanding debt in this way. For the 5,579 open market repurchases in the data between 1986 and 2013, the average percentage of the debt issue retired is 24.1%. Further, in only 7% of these cases was the company able to retire all of its debt. This mechanism, is, however, fairly inexpensive. The average price paid by the company to retire its debt in this way is 95.23% of par.

The second additional mechanism that the firm has for retiring early its debt is a tender offer, in which it submits a written offer to bondholders informing them of the firm's desire to retire its debt early, and offering a price (typically a premium) at which the firm can buy back its debt from bondholders. Across the 5,694 tender offers from 1986 to 2013, firms tended to pay a significant premium in tender offers, as might be expected from the fact that they are signaling to bondholders that they wish to retire the debt. Several studies, such as Mann and Powers (2007) have cited asymmetric information as one potential reason for these high premia. The median price as a percentage of par paid by firms was 107%, however, there is significant positive skew to the tender price distribution. In fact, the

75th percentile of the tender price distribution exceeds the 98th percentile of the call price distribution. The majority of these offers tend to be fixed price, according to Kruse et. al., but they can also take the forms of fixed spreads or Dutch auctions. Tender offers are also a somewhat more effective way for the firm to retire a significant portion of its debt than open market repurchases. Across our sample, the average amount retired was 62.5% of the total debt issue (other studies have found even higher figures for far more limited early samples), and, in 13% of the cases, the firm was able to retire the entire debt issue.[2]

In order to synthesize all of this information, it is helpful to present the same statistics for callability. Partial and complete calls constituted 92.78% of the early redemptions in our largest data sample. On average, the calls retired 92.2% of the outstanding debt issue, and 91.2% of calls retired the entire debt issue. The average price paid for all calls was 99.23% of par, and the average price paid for make-whole calls was 104.21% of par. Calls also retire far more of debt issues than either alternative method, are more often in the data, and are subject to significantly lower transactions cost and legal fees. Compared to open market repurchases, 91.2% of calls retire the entire bond issue, while only 7% of repurchases do. Moreover, firms that issue callable or make-whole debt still have the option to perform either open market repurchases or tender offers as firms with noncallable debt would. For some firms, the difference in early redemption price between callable debt and noncallable debt is not high. For many, however, it is. Figure 4 illustrates this by plotting the price paid for early retirement of debt as a percent of par for three types of bonds: noncallable, traditional callable, and make-whole. We see that while the series are fairly similar for the bottom 30% of prices, the upper 30% of prices paid to retire noncallable debt far exceeds that paid to retire either form of callable debt. In fact, the make-whole premium looks fairly small compared to the additional premium one might have to pay to retire noncallable debt. In summary, if a firm expects that there is some chance that it will refinance its debt early, both callable and make-whole debt seem to offer far less risky and more cost-effective ways

[2]A third potential mechanism for early retirement is a sinking fund provision, which can enable a firm to repurchase some of its debt each year, but these provisions are uncommon and somewhat limited in scope.

to do so.

2.3. **Data and Empirical Trends**

In this section we begin by describing the data sources employed and the methodologies used to trim the data. The second subsection then describes empirical results about the firms more likely to issue callable debt and the uses of the proceeds of such debt.

2.3.1. Data Sources and Methodologies

Two main sources of data were used for this project. The first is the Mergent FISD fixed income database, which provides bond information. This database contains several datasets that were useful for this project, among them the Bonds Issues Dataset, the Amount Outstanding Dataset, and the Redemption Dataset. The Bond Issues dataset contains bond-specific information for over 350,000 bond offerings between 1986 and 2014. In particular, this dataset was used to gather information such as bond par values, yields, maturity, coupons, issue dates, callability, and other option. The dataset also provides identifying information about bond and its issues, such as the issue CUSIP, issuer CUSIP, the industry of the issuer, and FISD-specific identification codes for both bond and issuer. One variable not included in this dataset is whether a bond is make-whole, and for this the FISD redemption dataset was used. After merging the Redemption dataset with the Bond Issues dataset, 193,776 observations remained. These observations were filtered to focus on U.S. corporate bonds issued by nonfinancial firms, and were then filtered to exclude certain uncommon options and features, such as fungibility, convertibility, lease obligation issues, etc., resulting in a final dataset of 20,166 bonds.[3]

For each of these bonds, FISD's Amount Outstanding dataset provides detailed descriptions of the instances where the amount outstanding of each debt issue potentially changed. It begins with the issuance of the debt and ends with the maturity or early retirement of the

[3]More details about the exact bond features excluded and included are available upon request.

debt, and seems to be the most complete source of bond calls, open market repurchases, and tender offers. For each action, the dataset identifies the relevant issue, the type and date of the action, and the amount outstanding both before and after the transaction, as well as the price of the transaction, expressed as a percentage of the issue's par value. For the data in our sample, this dataset contained between 2 and 8 actions for each bond issue. We merged this and the other FISD datasets using the bond-specific issue id, which uniquely identifies each bond issue. The final dataset has just over 100,000 action-level observations.

This bond data was supplemented with the second main source of data for this project: firm data from Compustat. For this we used the Quarterly and Annual Fundamentals datasets to obtain over 50 firm-specific variables, primarily balance sheet and income statement data. Among the most important of these were measures of assets, debt, and equity levels, debt flows, dividend policies, investment flows, M&A activity, interest payments, and revenue/net income figures. We then merged this Compustat data with the combined FISD data by matching either the CUSIP values, the company tickers, or the first five CUSIP digits and the company ticker. After conducting all of these merges, we then compared the results across merge categories, finding no significant differences in variables after using a relevant Holm correction for our .05 alpha level and the number of pairwise tests.

2.3.2. Empirical Trends

The purpose of this section is to establish two stylized facts that will be important in motivating the model in Section 5. The first fact is that firms with more credit risk are more likely to use callable debt (of all forms.) Second, we show that the usage of callable debt is closely tied to firm investment policies. Specifically, callable debt (and in particular make-whole callable debt) is more likely to be issued to fund future investment. We show these two facts by examining both the types of firms issuing callable debt and the use of the proceeds from debt issuance.

We begin by examining the cross-sectional characteristics of the firms issuing callable debt.

Since over 90% of bonds in our sample in recent years have been callable, we consider the trends in callability over time of firms with different characteristics. Using this, we can draw inferences from both the rate of adoption of call provisions and the overall level of the prevalence of callable bonds.

The first cross-sectional characteristic that we consider is a firm's credit risk. There are many different measures that one could use for this, but we choose the S&P Long-term Issuer Credit Rating as a reasonable summary statistic of all of the factors that impact the creditworthiness of a company. Similar results also hold if one uses leverage or the volatility of earnings. For this analysis, we use the merged Compustat-FISD datasets and sort firms based on their credit ratings. We then separate bond issuances based on the credit ratings of the firm issuing them at the time of issuance, and consider how the trends in callability have varied across these firm credit ratings. Figure 5 displays the results. We see quite clearly that firms with lower credit ratings have always had a higher level of callability in their bonds, and that they were faster to adopt call provisions, and that this trend holds across all three groups considered. In summary, it seems evident that credit risk and the inclusion of a call provision are positively related: the higher a firm's credit risk, the more likely it is to include a call provision in its bond.

The second fact that we hope to show is that investment and the inclusion of call provisions are directly linked. We do this by showing two relationships. First, firms that invest more tend to issue more callable debt. Second, firms that issue callable bonds (in particular make-whole callable bonds) are more likely to use the proceeds of their bond issuance for investment.

We begin by considering again the trends in callability across different levels of firm investment. We measure investment here by firm capital expenditures[4] and perform a similar sorting exercise to that done previously. Namely, we pair bonds to matched annual capital

[4]R&D/intangible capital not included due to a lack of reliable data

expenditure to operating income ratios for the firms issuing those bonds, and then sort those observations into quartiles based on the ratio in each year (so as not to pick up the effect of average ratios changing over time). We then plot the trends in the prevalence of call provisions for each investment quartile in Figure 6. We see that firms in the lowest quartile of investment issue a lower fraction of their bonds as callable: in recent years this level has been roughly 20% less. Furthermore, it seems that these firms are more responsive to market conditions in their callability. While firms with higher levels of investment maintain a high fraction of callable bonds across market conditions, firms with lower levels of investment are more likely to decrease their usage of call provisions in bad times, such as the recent financial crisis.

The final form of evidence for the link between investment and the issuance of callable bonds comes from studying how firms use the proceeds of bond issuance. For this we consider a slightly different test. Since our outcome variable is now continuous on a firm-by-firm basis, we perform a fixed-effects regression of the post-issuance level of firm accounting variables on their pre-issuance level and a dummy for whether the bond issue was a make-whole callable bond. (We can use make-whole callable bonds now since we are not merely considering the trend over a few years.) Table 1 displays the results. We see that, relative to non-make-whole bond issues, issuers of make whole bonds use less of the proceeds for cash and dividends and far more for investment into property, plant, and equipment. This again demonstrates the link between callable bonds and investment: firms that issue callable bonds, in particular make-whole callable bonds, are those firms who have tended to invest a higher fraction of their income and who tend to use more of their proceeds for investment.

Thus, this section has demonstrated two empirical facts. First, firms with lower credit ratings (more credit risk) are more likely to issue callable bonds. Second, firms that invest more and that are more prone to use their bond proceeds for investment are more likely to issue callable debt. These facts will help motivate the model in Section 5.

2.4. **Analysis of Previous Explanations**

Before beginning with our model, we briefly explore previous explanations for the issuance of callable debt and show why these explanations cannot rationalize the current trend. We explore four theories: asymmetric information, risk shifting, underinvestment, and interest rate risk management.

2.4.1. Asymmetric Information

The first theory concerns asymmetric information. As suggested in Barnea, Haugen, and Senbet (1980) (BHS), managers may have more information than is available to public investors (in particular bondholders.) If this is the case, managers with positive private information who issue non-callable bonds prior to the revelation of the information will be sharing the surplus of the revelation of that information with bondholders. This is because the revelation of that positive information will presumably reduce the default risk of the firm (or more generally improve its creditworthiness), increasing the value of its bonds. Existing bondholders will realize all of this benefit while managers and equityholders will not benefit from the appreciation in value of the bonds.

BHS suggest a solution to this problem in the form of callable debt. Since the call option is held by equityholders will appreciate in value by the same magnitude as the bonds, this security will appropriately compensate equityholders for the revelation of positive information. Just as bondholders undervalue the firm's creditworthiness (relative to the full information case), they also undervalue the call option by the same amount and so equityholders are appropriately compensated.

While this may be true for an at-the money call option where the delta of the option is approximately one, it is certainly not the case for make-whole debt, and therein lies the issue when one applies this theory to the current trend. As stated previously, the options on make-whole bonds are structured so that they are almost never in the money. Since the strike price contains such a low spread to the benchmark Treasury, the firm's credit profile would have to improve enormously for the value of the underlying bond to exceed the strike price. This means that the option is initially deeply out of the money, and, as such, has a delta far below one. So, even if the firm's credit profile were to improve, the value of the option would not increase by the same value of the bond, and, in fact, would hardly increase at all. Since these options are virtually never in the money, their value would remain close to zero over the life of the bond and even revelations of positive information are not likely to change that. As such, equityholders will receive very little compensation for their private information.

What this implies is that equityholders with private information will be poorly served by seeking to mitigate this wealth transfer by issuing make-whole callable debt. There are a number of alternate solutions, including shorter maturity debt, convertible debt, and bonds with call options that are not make-whole. But clearly, the increase in the issue of make-whole bonds cannot be rationalized by managers seeking to ensure that equityholders are compensated for their private information.

2.4.2. Risk Shifting

A second issue identified by BHS that may motivate the issue of callable bonds is risk shifting. The idea is that after issuing noncallable debt, equityholders' claims on the firm's assets will be subordinate to a higher fixed claim by bondholders. The "call option" that equityholders hold on the value of the firm has a higher strike price as more debt is issued. In maximizing the value of the firm to equityholders, therefore, managers may be incentivized to take on riskier projects. If debtholders expect this action ex-ante, then it will naturally reduce the price that they pay for debt when it is issued. The conflict here comes from the fact that taking on such risky projects reduces value for bondholders while increasing it for equityholders. One potential solution to this proposed by BHS, then, is to issue callable debt. Since the bond value decreases with the adoption of these projects, the value of the call options held by the equityholders will decrease, and this will act to temper any incentives that equityholders have to take on these projects.

In discussing this solution, Crabbe and Helwege identify a key element to eliminating the conflict: "To eliminate the incentive to increase risk, a firm will include a call option whose value equals the potential gain from switching investments" (page 5). While this may be possible for non-make-whole callable debt, it is certainly not possible given the typical structure of make-whole issues, for similar reasons as discussed above. In particular, since the call options in make-whole issues are almost always significantly out of the money, the value of those options is not likely to change much based on firm investments. Since the value starts out very low for the vast majority of these issues, the adoption of risky, low-NPV projects by the firm cannot decrease the value by much and so will not act as an effective counterweight to the incentives for equityholders to increase the riskiness of the firm's value.

Clearly then, this motive cannot explain the recent increase in make-whole debt. Additionally, as Crabbe and Helwege observe, this theory would also imply that riskier firms should issue bonds of lower maturities, but this appears empirically to not be the case. For example, Xu (2016) shows that the average maturity of speculative-grade bond issues is significantly lower than that of investment-grade bond issues.

2.4.3. Underinvestment

The third agency theory that we consider is the underinvestment problem proposed by Myers (1977). This problem arises when managers, after issuing debt, receive an investment opportunity that is likely to only provide a payoff to bondholders. An example of this, discussed in Bodie and Taggert (1978) is a firm that has nontrivial default risk receiving news of a fairly safe project that provides a fairly low payoff. Managers seeking to maximize shareholder wealth would then prefer not to make this investment, saving their capital for projects that can potentially benefit equityholders. As with the previous case, this will reduce the ex-ante price paid for the debt by bondholders.

The solution proposed by Bodie and Taggart is the embedding of call options in these bonds. The call options alleviate this problem by allowing the firm to recontract based on the new project/investment and thus allowing equityholders to be compensated for the adoption of this project. While this may hold for traditional callable bonds, it fails with make-whole bonds. This is because even with the revelation of a new investment project beneficial to bondholders, the option embedded in a make-whole bond is still highly unlikely to be in the money. The spreads to benchmark Treasuries that characterize the strike price of the option are so low that the bond would need to be extremely close to risk-free for this to occur. Given that this problem is most acute for firms with significant default probabilities, such a transformation is wholly unlikely. Furthermore, given that there are several other mechanisms by which firms can mitigate this conflict (for example shorter maturity bonds), this hypothesis also cannot explain the risk of make-whole bonds.

2.4.4. Interest Rate Risk Management

The last explanation for the usage of callable debt that we consider is interest rate risk management. Although proponents of this justification have decreased over time, it is simple to show that make-whole bonds do not offer any interest rate hedges for firms. The idea behind this hypothesis is that firms that issue bonds at a high interest rate may seek to refinance at a lower rate, benefiting from the reduction in interest payments. Bonds with traditional call options enable firms to do this by allowing them to repurchase their debt at its par value and issue debt at the prevailing (lower) market interest rate. Kraus (1973) showed that this is a zero-sum game and hence should be priced equally by both parties (assuming the same stochastic discount factor) and the widespread use of OTC derivatives by corporations seems to have eliminated the need for this.

Even more strongly, note that make-whole bonds do not help firms manage interest rates. Since the call price paid by the firm reflects the prevailing market interest rate, firms that issue bonds at high interest rates and seek to refinance at low interest rates will be forced to pay at least the market price of the bond, and thus will at best earn no profits. The graph and subsequent explanation in Section 2.1 illustrate this point. Clearly then, the motive of hedging interest rate movements cannot be behind the increase in make-whole bonds.

We have thus seen that several of the most popular explanations for why firms issue callable bonds fail to explain the recent increase in make-whole callable debt. The asymmetric information and underinvestment hypotheses both require managers to be potentially interested in exercising the call option to recontract, something which is highly unlikely given the structure of make-whole call options. The risk shifting motive requires the option to be priced such that its sensitivity to price decreases in the underlying bond is relatively high, which is again improbable since the option is deeply out of the money at issuance. The predictions of this theory regarding the interaction of cash flow riskiness and debt maturity also seem to contradict recent empirical evidence. Finally, the interest rate risk management story cannot rationalize make-whole bonds almost by construction: the make-whole option is designed to insure bondholders against the risk of changes in market interest rates, not firms.

2.5. **Model and Results**

It thus seems that we need a new theory to explain why firms have been increasingly issuing make-whole debt. We propose that theory in this section, beginning by motivating it and providing some background in the first subsection. The second subsection explains the mechanism of the model and the third presents results of the discrete model. We extend this model to a fully dynamic infinite-horizon model in section four.

2.5.1. Motivation

The main idea behind our model is refinancing risk. Put simply, firms realize that at the time at which they refinance their debt, the availability of credit and the credit spread they pay is determined based on credit market conditions at that time. If firms wait until their debt matures to reissue, then they are forced to either be subject to the prevailing credit market conditions or seek other forms of financing, both of which can be costly. In seeking to avoid these costs, firms may prefer to have a choice of refinancing dates on or before the maturity of their debt issues. Make-whole debt allows firms to do this. This hypothesis is consistent not only with the empirical trends that we documented, but also with the results that others have found–both theoretical and empirical. We begin with a bit of background of refinancing risk.

There has been significant work showing that refinancing during tight credit markets can be costly to firms. Firms may have to refinance at significantly higher interest rates (Froot et. al. 1993) or worse bond terms (He and Xiong 2012). If financing is in short supply or altogether unavailable, firms may be forced to liquidate excessively by creditors (Diamond 1991), sell assets in a firm sale (Choi et. al. 2013) or decrease investment (Almeida 2009). Of course, firms may also be forced to default (He and Xiong 2012). Moreover, this seems to be a concern that both firms and financiers recognize. Graham and Harvey (2001) show that CFOs claim they manage debt maturity to "reduce risk of having to borrow in bad times," while credit rating agencies commonly cite refinancing risk as a reason to downgrade firms (and the refinancing of debt as a reason to upgrade firms.) A concrete example can be found in Bank of America Merrill Lynch's 2012 advice to CFOs:

"Don't wait too long to refinance upcoming maturities. Give yourself at least 18 months before your current financing matures, so that if any segment of the market ... shuts down for a few months, you'll still have time to get something done when the markets inevitably return to life"

Bank of America cites that advice as being one of the "lessons from the financial crisis", and indeed this is a risk that is naturally heightened by financial crises and observed credit market freezes. In the model we will tie this refinancing risk to the issuance of make-whole debt, and we note that the fact that the issuance of make-whole debt began to increase significantly during the financial crisis is one piece of evidence that this link is valid. Another comes in the cross-sectional characteristics of the make-whole bond issuers. In Section 3.2 we observed that firms with higher credit risk are those that issue make-whole debt at a higher rate, and these are precisely the firms that are more likely to be affected by refinancing risk, since their credit spreads are more countercyclical and their probability of accessing the credit market more procyclical. In addition, these firms tend to be more constrained in the debt maturities that they can issue, as Xu (2016) shows. The other trend that we observed in Section 3.2 is that the issuance of make-whole debt and your investment policies are closely linked: firms that invest more of their earnings tend to be more likely to issue make-whole debt. This too is evidence for the refinancing risk explanation: refinancing risk is highest when a firm requires a steady stream of income. If the firm can vary its assets side with variations in its liabilities, refinancing risk is not as large of a concern. Of course, capital investment is one of the more irreversible forms of capital (see for instance Ramey and Shapiro (2001)). Thus, firms with heavier investment policies are likely to have assets that are harder to adjust downwards in level and thus are likely to be more sensitive to refinancing risk.

It is also worth mentioning the several papers that have been devoted to showing this refinancing risk directly. These include Mian and Santos (2011) and Julio (2013), but perhaps the most relevant is Xu (2016). Xu shows that speculative grade firms "frequently refinance early to extend the maturity of their outstanding bonds, particularly under accomodating credit supply conditions" and concludes that "the evidence is consistent with precautionary maturity management, in which speculative-grade firms extend maturity to hedge against refinancing risk caused by credit supply fluctuations."

2.5.2. Model Setup

We now propose a simple model to capture this refinancing risk effect and examine the impact of (make-whole) callable debt. The model features four periods and firms who invest in the first period. The firms finance this investment through their own initial equity and by issuing debt of one of two forms: non-callable or make-whole. That is, they cannot raise additional equity and we do not consider the choice between traditional callable and make-whole. Firms are subject to idiosyncratic productivity shocks which determine the return from their investments. The key friction in the model comes in the timing of investment and financing: firms are constrained to issue at most two-period debt, but their investment takes either two or three periods to mature. Liquidating investment before maturity is inefficient, so firms are incentivized to refinance their debt if their investment takes three periods. In this case, a firm that issued traditional callable debt has to refinance at maturity of its debt (period 2), whereas a firm that issued make-whole debt can refinance in either periods 1 or 2. We will see that his expands access to credit markets and affects the optimal capital choice for these firms.

Firms begin with an initial amount of equity E_0 that is held in cash and an opportunity at the initial time period to invest capital k_0 in a technology yielding $y_t = Z_t k_0 - \alpha k_0^2$. Z_t represents the firm's idiosyncratic productivity, and its log follows an AR(1) process:

$$z_{t+1} = (1-\rho)\mu_z + \rho z_t + \sigma_z \epsilon_{t+1}$$

where ϵ_{t+1} is drawn from a standard normal distribution. The returns from this technology materialize at t=3 with probability p and at t=2 with probability $1-p$. To finance this capital k_0, firms issue 2-period debt D_0 at time 0 maturing in time 2. They have two choices for this debt: non-callable debt, which must be refinanced in period 2, and make-whole debt, which can be refinanced in either period 1 or period 2. We denote the respective interest

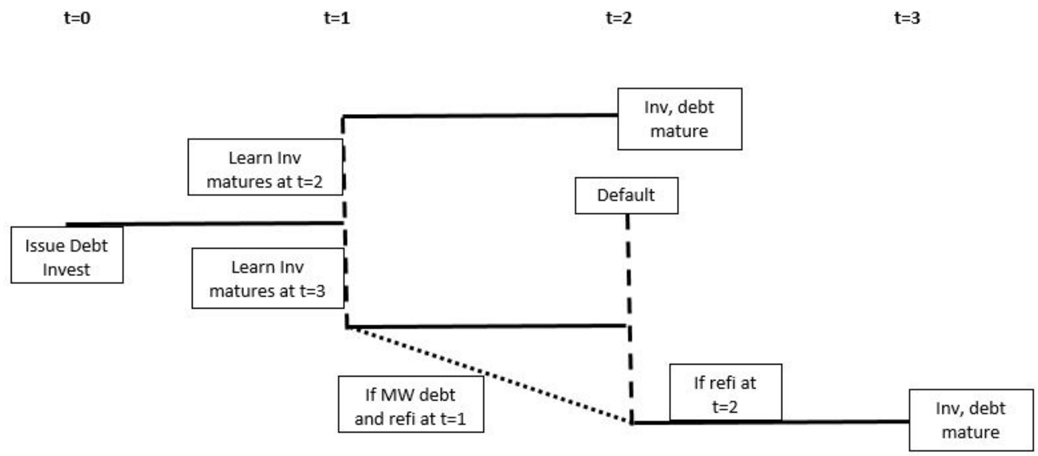

Figure 10: Make-Whole Model Timeline

rates for non-callable and make-whole bonds as r_0^{NC} and r_0^{MW}. Debt is fairly priced, and both firms and bondholders use a constant discount factor β.

The firm learns in period 1 whether its investment returns will be realized in period 2 or period 3. If the investment matures in period 2, then the initial maturity of its debt and the maturity of its investment match, and the firm does not need to refinance its debt. However, if the investment matures in period 3, the firm will need to obtain funding between periods 2 and 3, which requires it to refinance its debt. If the firm is unable to do so, it is forced to default and its salvage value becomes ψk_0. The following diagram illustrates the timing of events in the model:

The following diagram illustrates the timing of events in the model:

Note that the firm has no incentive to issue make-whole debt when it refinances since the flexibility in refinancing date is of no value to it once it knows its investment maturity and can issue debt with the same maturity. Thus we can denote the interest rate at which the firm refinances as r_t^{NC} where either t=1 or t=2 depending on the date of refinancing. Also note that the firm prefers to use its costless initial equity first, then finance any further investment with debt. Thus, given a level of investment, the amount that it borrows can be written as $D_0 = max\left\{0, k_0 - E_0\right\}$. Given this, we can write the firm's value if it issues non-callable debt as follows:

$$V_0^{NC} =$$
$$\underset{k_0}{max} E\left[p\beta^3 max\left\{Zk_0 - \alpha k_0^2 - D_0\left(1 + r_0^{NC}\right)^2\left(1 + r_2^{NC}\right), 0\right\} 1_{\{refi\}} + (1-p)\,\beta^2 max\left\{Z_2 k_0 - \alpha k_0^2 - D_0\left(1 + r_0^{NC}\right)^2, 0\right\}\right]$$

where the first term represents the value if the investment takes three periods to mature and the firm is able to refinance and the second term represents the value if the investment takes two periods to mature. Similarly, the firm's value if it issues make-whole debt is given by:

$$\begin{aligned} V_0^{MW} = \underset{k_0}{max} E[p\beta^3 \Big(& max\left\{Y(k_0) - D_0\left(1 + r_0^{MW}\right)^2\left(1 + r_2^{NC}\right) - MWprem, 0\right\} 1_{\{refi@t=2\}} + \\ & max\left\{Y(k_0) - D_0\left(1 + r_0^{MW}\right)\left(1 + r_2^{NC}\right)^2 - MWprem, 0\right\} 1_{\{refi@t=1\}} + \\ & (1-p)\,\beta^2 max\left\{Y(k_0) - D_0\left(1 + r_0^{MW}\right)^2, 0\right\}] \end{aligned}$$

where the first term represents the value if the investment takes three periods to mature and the firm refinances in the second period, the second term represents the value if the investment takes three periods to mature and the firm refinances in the first period, and the third term represents the value if the investment takes two periods to mature (in which case the firm does not need to refinance its debt).

Thus the firm will optimally choose both its level of borrowing (and hence its level of investment) and its type of borrowing. We will see that the choice of the latter can be quite important in terms of the firm's access to credit markets and the price it pays for that access.

2.5.3. Results

There are two key results from this model. First, we show that having access to make-whole debt can increase a firm's access to capital. In particular, refinancing risk can lead to a firm being frozen out of time-0 credit markets if it attempts to issue non-callable debt. By allowing the firm to refinance in two different periods, make-whole debt reduces the refinancing risk faced by the firm and so can alleviate market shutdowns in time 0. Second, make-whole debt increases access to capital: firms can generally issue more debt as make-whole than otherwise.

We demonstrate these effects by considering whether a firm with a given set of parameters will be able to access credit markets for each type of debt. The outcome variable here is how open or closed the credit market is, which is measured by whether the Euler equation for bondholders can be satisfied for some interest rate and, if not, the minimal gap across interest rates. Thus a value of zero corresponds to open credit markets, while larger figures correspond to credit markets that are farther from being open. We first fix the underlying productivity process and consider how access to credit markets varies with a firm's initial productivity state (on a scale of 1-21) and the amount of capital the firm is seeking to invest. The following plots illustrate the credit market outcomes[5]:

The top plot concerns non-callable debt while the bottom plot presents results for make-whole debt. For each plot, the x-axis represents the firm's initial productivity level (on a 21-point grid). Higher numbers here correspond to higher initial productivity levels (which correspond, in turn, to reduced credit risk.) The y-axis plots the level of investment, k_0, that the firm wishes to undertake. This investment is directly linked to the amount of debt that firms issue since they will fund investment first with their limited initial equity, and

[5] These results based on $\mu_z = .04$, $\rho_z = .5$, $\sigma_z = .025$, $E_0 = .1$, $\beta = .99$, $\alpha = .2$, $rr = .6$, $p = 1$, $MWprem = .5 * \left(r_0^{MW} - \left(\frac{1}{\beta} - 1\right)\right)$.

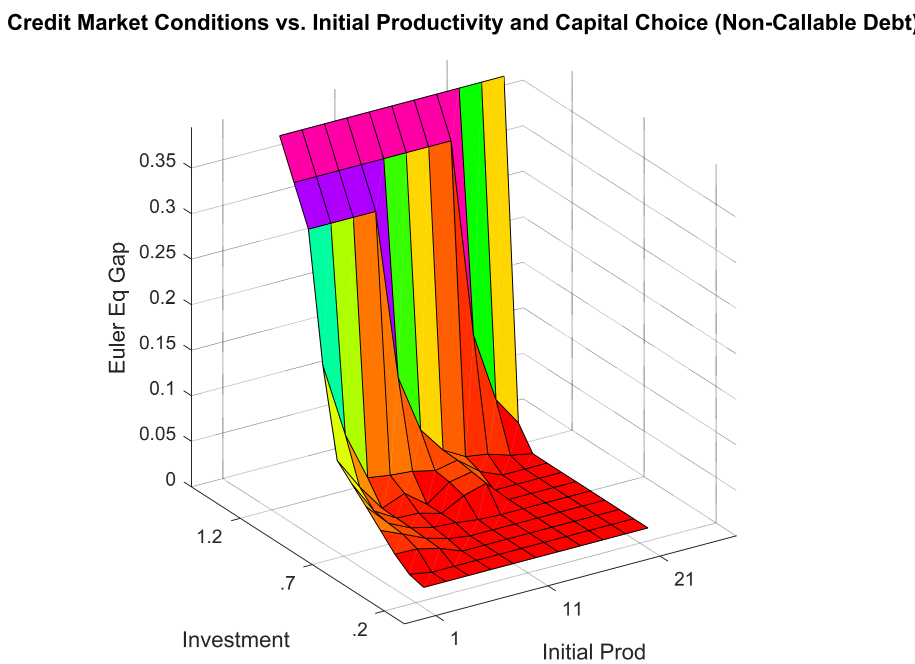

Figure 11: Euler Equation Gap for Non-Callable Bond

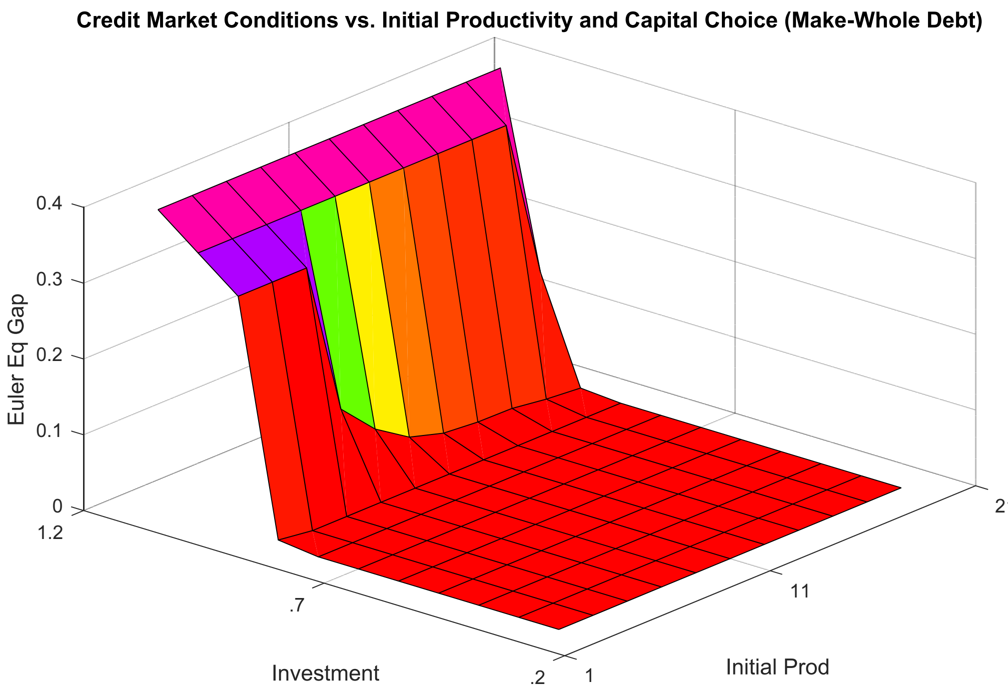

Figure 12: Euler Equation Gap for Make-Whole Bond

then by issuing debt. Lastly, the z-axis gives a measure of how open or closed credit markets are. Specifically, it plots the minimal gap in the Euler equation across all interest rates for debt of a specific amount and for a specific firm. Levels of zero correspond to the firm being able to access credit markets and borrow that amount, while levels above zero imply that there is no interest rate satisfying the bondholder's Euler equation. In the latter case, the magnitude of the variable on the z-axis indicates just how significant the credit shutdown is: it gives a measure of the dollar transfer that the lender would need to make the loan.

Thus we see that access to make-whole bonds has two significant effects. First, for a given level of investment, poorer-quality firms gain access to credit by issuing make-whole bonds whereas they otherwise would not be able to access credit markets through non-callable debt. For investment levels near the middle of the distribution, only about the top half of firms (in terms of initial productivity) have access to credit market through non-callable bonds, whereas all firms can access credit markets through make-whole bonds. We see that this effect is greater for firms with lower initial productivity levels, corresponding to the empirical pattern that firms with lower credit ratings are more likely to issue make-whole debt. Second, for a given level of initial productivity, firms can borrow far more with make-whole debt than they would be able to with non-callable debt. For firms in the middle of the productivity distribution, make-whole bonds allow them to borrow roughly twice as much as non-callable bonds. This also matches the empirical evidence that issuance of make-whole debt and high-investment policies tend to be significantly linked.

2.5.4. Full Model

We now extend this model to an infinite-horizon setting in which firms dynamically choose their refinancing policy. Investment, as in the previous model, takes place entirely in the first period. Firms finance their investment first with initial equity and then with one of four instruments: one-period debt, two-period noncallable debt, two-period make-whole debt, and additional equity. The projects have stochastic maturity where the completion date of the project follows an exponential distribution with parameter λ. Firms need to maintain their initial source of funding until their project matures, at which point they realize cash flows from the project, pay financing costs, and distribute the rest of the proceeds to initial equityholders. As before, firms are subject to idiosyncratic productivity shocks and both equityholders and bondholders discount cash flows at a constant discount factor β.

Note as before that refinancing risk here comes from both the stochastic maturity of the project and the time-varying firm conditions. The firm is required to maintain its source of financing until a realization of the project maturity, but in the meanwhile its idiosyncratic productivity independently fluctuates. Moreover, as time passes, interest accumulates on debt that the firm has borrowed, thus requiring the firm to finance increasingly large amounts to continue the project. (This is one significant sense in which there is time dependence in this problem and it cannot be thought of as a series of static problems; another is the autocorrelation of the productivity shock.)

Comparing the forms of financing, we see that they have very different implications for refinancing risk and firm value. Additional equity never needs to be refinanced, but is subject to issuance costs and does not feature the interest rate tax shield. One period debt must be refinanced every period and two period noncallable debt must be refinanced every other period (in the absence of a project maturity.) Two period make-whole debt may be refinanced either one or two periods after issuance. The firm will optimally choose to refinance one period after issuance if credit markets are open (i.e. there is no gap in the Euler condition) and equity value is maximized compared to waiting. It is important to keep in mind that it is not always optimal for the firm to refinance in the period following issuance as it will likely pay a higher interest rate (it will be refinancing at a higher leverage).

The payoffs in each state largely follow from the previous model, and the time-0 value functions can be written in simplified form as follows (the time-t value functions are identical but for the optimization over the capital stock):

For a firm issuing one-period debt:

$$V_0^{one} = \max_{k_0} E\left[\lambda\beta(\text{payoff from project maturing next pd}) + (1-\lambda)\beta 1_{\{refi\}} V_1^{one}\right]$$

For a firm issuing two-period noncallable debt:

$$V_0^{two,nc} = \max_{k_0} E\left[\lambda\beta(\text{payoff from project maturing next pd}) + \right.$$
$$\left. \lambda(1-\lambda)\beta^2(\text{payoff from project maturing in two pds}) + (1-\lambda)^2\beta^2 1_{\{refi\}} V_2^{two,nc}\right]$$

For a firm issuing two-period make-whole debt:

$$V_0^{two,mw} = \max_{k_0} E\left[\lambda\beta(\text{project payoff next pd}) + (1-\lambda)\beta 1_{\{\text{refi in 1}\}} V_1^{two,mw} + \right.$$
$$\left. \lambda(1-\lambda)\left(1 - 1_{\{\text{refi in 1}\}}\right)\beta^2(\text{project payoff two pds}) + (1-\lambda)^2\beta^2 1_{\{\text{refi in 2}\}} V_2^{two,mw}\right]$$

For a firm issuing additional equity:

$$V_0^{e} = \max_{k_0} E\left[\lambda\beta(\text{payoff from project maturing next pd}) + (1-\lambda)\beta V_1^{e}\right]$$

We utilize this framework to ask two major questions. First, what are the investment impacts of firms having access to make-whole debt? That is, do firms invest more when they can finance this investment with an instrument that mitigates refinancing risk? Second, how much of an impact does this additional instrument have on the overall equity value of a firm?

For the first question, we see that firms that issue make-whole debt often invest significantly more than those issuing other forms of financing, and that this effect is stronger for firms

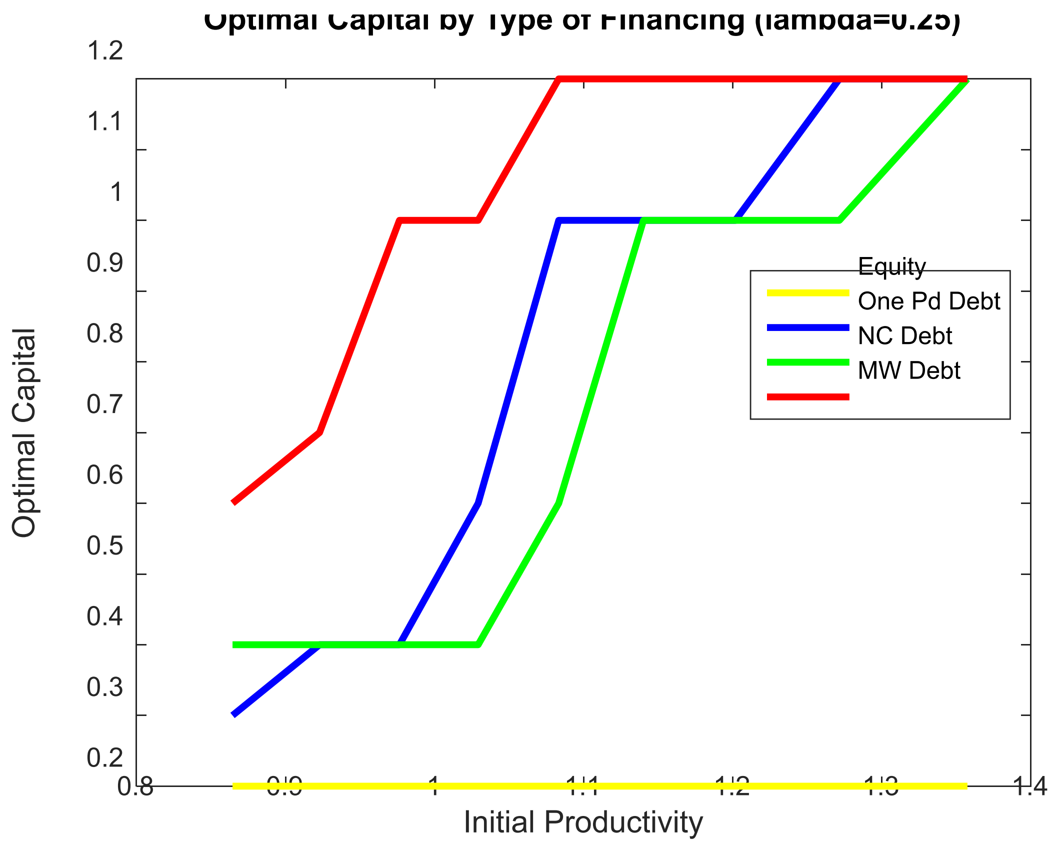

Figure 13: Optimal Capital Choice for Firms with Longer Projects

that face greater refinancing risk (in the form of longer project maturities. The two figures below illustrate this effect (recall that a higher lambda means that a project is more likely to mature sooner). For both this and the following sets of plots, the results are shown for time-0 firms (and hence leverage is computed using the initial equity and desired capital levels). As this leverage increases, the refinancing risk increases and the effects shown are more significant.

The differences in the level of investment are often stark, with firms investing 50-100% more in certain cases with make-whole debt compared to with other forms of debt. We similarly see that the differences in equity values vary with the level of refinancing risk (again through the channel of lambda):

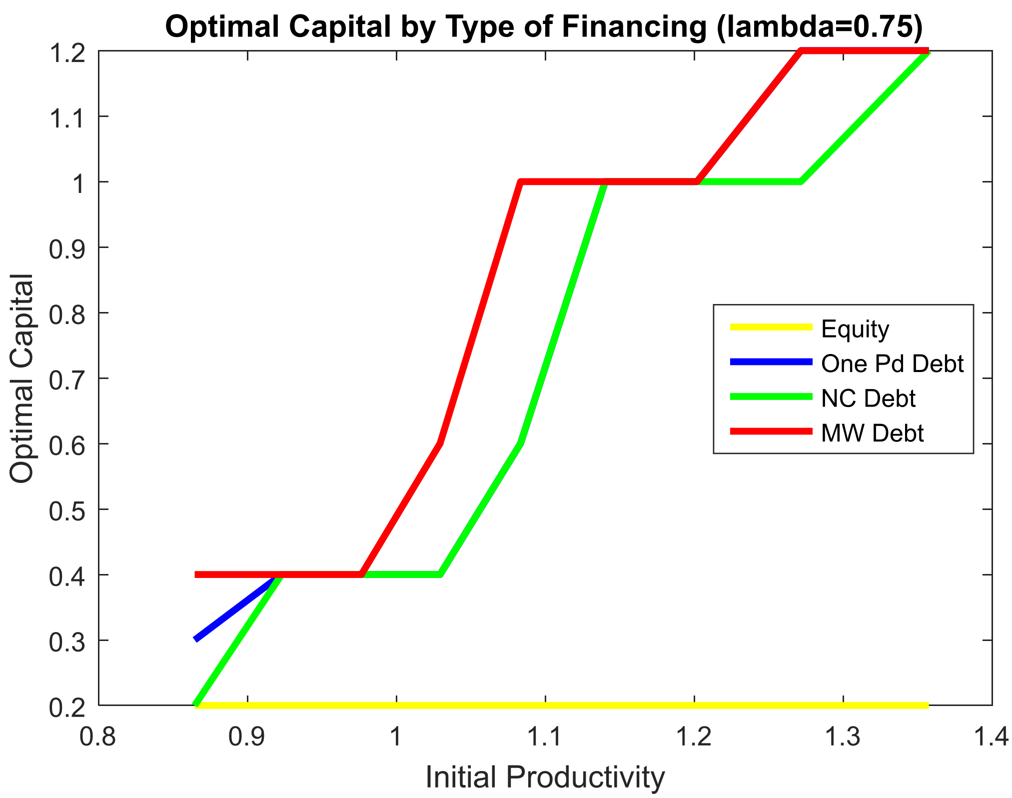

Figure 14: Optimal Capital Choice for Firms with Shorter Projects

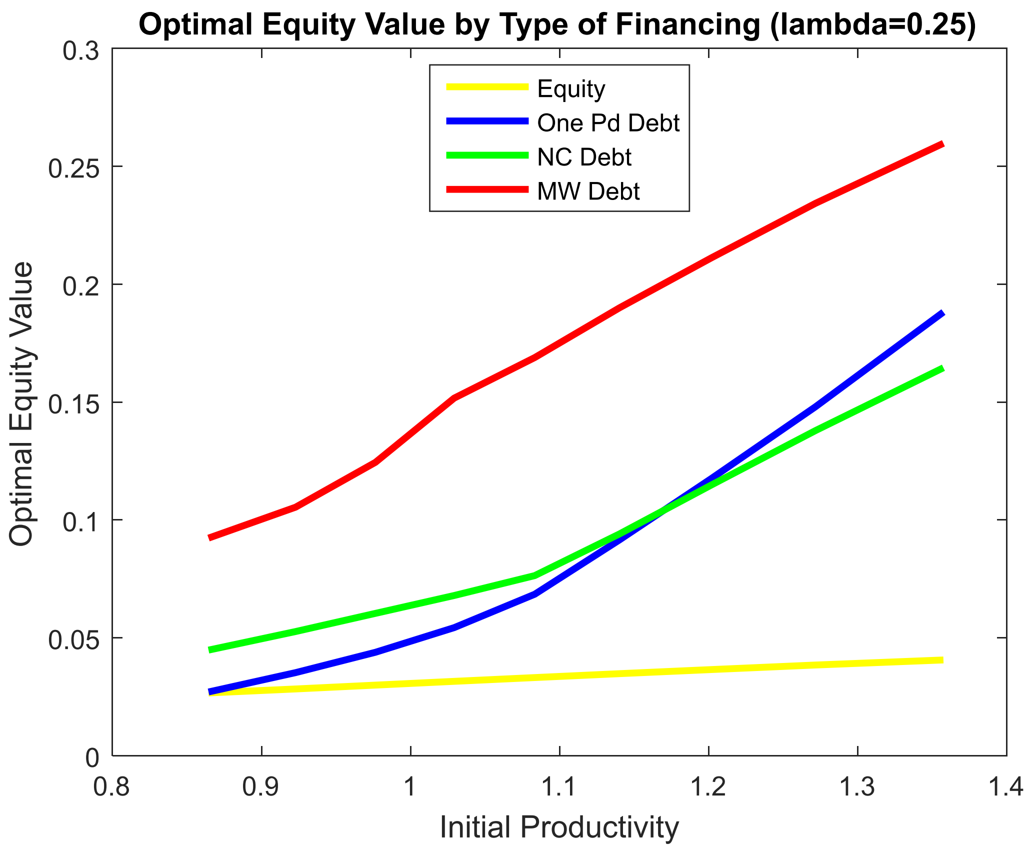

Figure 15: Equity Value for Firms with Longer Projects

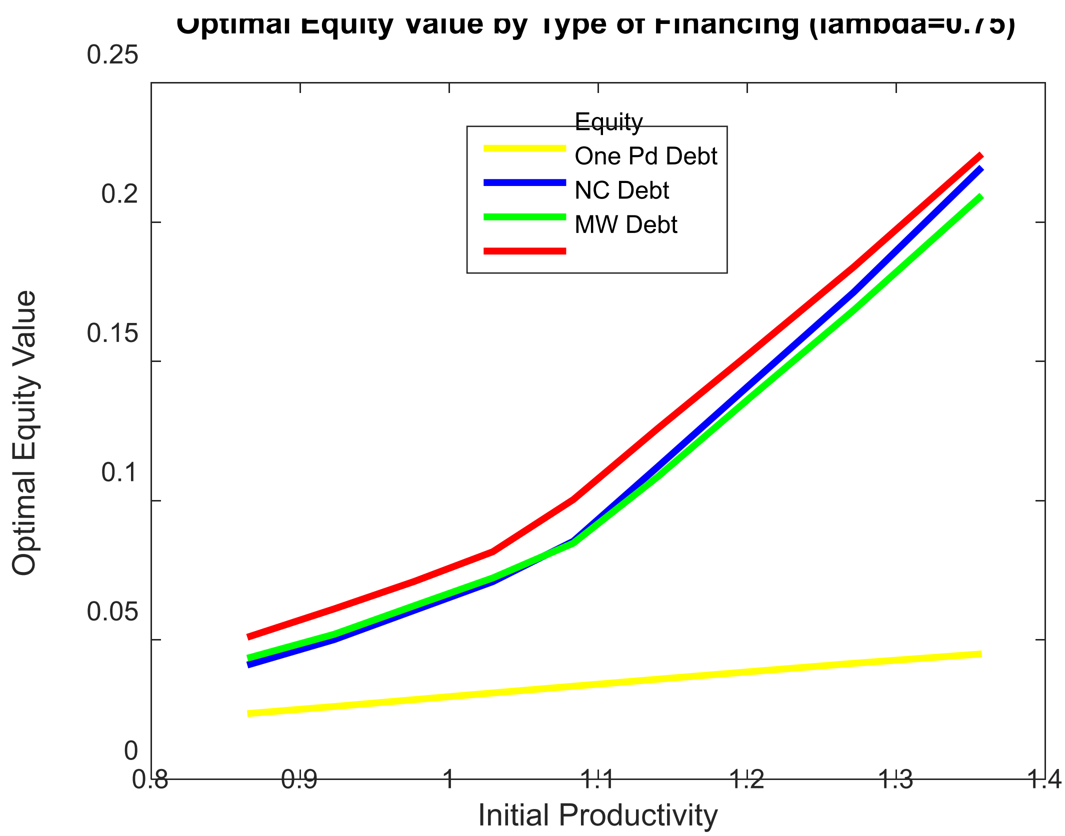

Figure 16: Equity Value for Firms with Shorter Projects

We see that the differences in equity value can also be significant. Firms issuing make-whole debt have equity values that are 4.9-11.2% higher in the high lambda case, and 27.6-51.5% higher in the low lambda (higher refinancing risk) case. Consistent with the empirical evidence, we also see that, the lower a firm's idiosyncratic productivity (and thus the higher its credit risk), the greater its benefits from issuing make-whole debt.

2.6. **Conclusion**

This paper began by demonstrating that not all callable bonds are created equal, and that the differences among callable bonds are in fact very important. We saw that call provisions have become exceedingly popular in bonds issued by U.S. nonfinancial corporations, but that much of this recent increase came from make-whole bonds. These make-whole bonds differ from traditional callable bonds in one very important way: their embedded call options are set at a price so low that they are virtually never in the money. This had several implications. First, we saw that previous motivations for the issuance of callable bonds based on mitigating asymmetric information and underinvestment problems by giving shareholders an option that would appreciate with the value of the bond no longer hold. Since the value of the make-whole call option does not move significantly with the value of the underlying bond, the incentives to equityholders are critically weakened. Second, the mitigation of risk shifting as an explanation of callable bonds was no longer valid either. The decrease in option value from taking on risky projects pales in comparison to the potential benefit to equityholders and again the mitigation of incentives for managers is simply not strong enough. Lastly, make-whole bonds do not allow firms to engage in interest rate risk management by construction: the price the firm pays varies with market interest rates.

Given these issues with existing explanations for the issuance of callable bonds, we sought to propose our own rationale. In order to do this, we first established three empirical facts.

First, the issuance of make-whole bonds began to increase significantly around the onset of the financial crisis in the U.S. Second, the issuance of make-whole bonds was (and is) far more prevalent for lower-rated corporations than for higher-rated ones. Third, investment policies and the issuance of make-whole bonds are closely linked: firms that invest more (as a fraction of their earnings) are more likely to issue make-whole bonds, and firms that issue make-whole bonds tend to invest more of the proceeds.

These empirical facts motivated the use of refinancing risk as a mechanism for explaining the issuance of make-whole debt. In particular, firms face the risk of rolling over their debt in tight credit market conditions and as a result prefer flexibility in when they can refinance their debt, something afforded to them by make-whole bonds. We showed that this is a stronger motive for firms with higher credit risk since they are more sensitive to credit market fluctuations and for firms that invest more since their capital is more irreversible. Then we embedded this friction into a simple model and showed that make-whole debt achieves two powerful benefits. First, it expands access to credit markets for firms with lower levels of productivity, and, second, it allows almost all firms to borrow and invest more. These features help explain why make-whole bonds have become so common and why their increasing prevalence in response to increased awareness of the effects of tight credit markets makes sense from a firm's perspective.

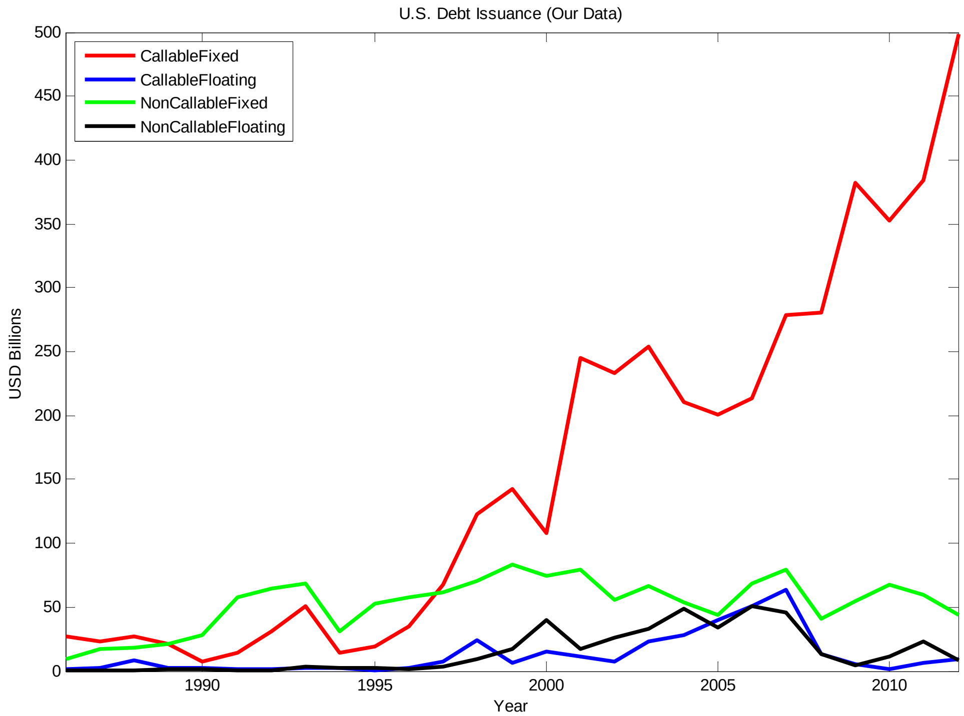

Figure 17: Total Par Value of Debt Issuance
After decreasing in popularity in the early 1990s, callable bonds have become far more common and currently represent the overwhelming majority of bonds issues by U.S. nonfinancial corporations.

2.7. **Appendices and Figures**

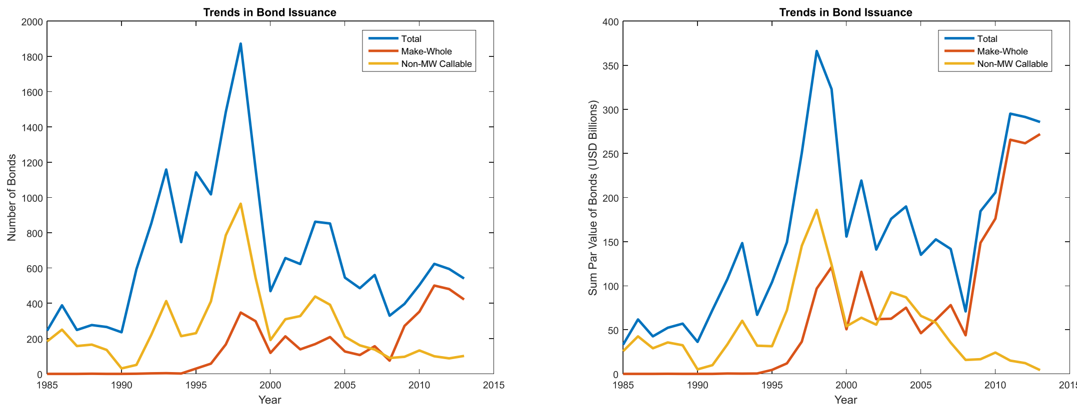

Figure 18: Total Debt Issuance by Type of Callable Bond
The usage of non-make-whole callable bonds decreased significantly post-1999, while the usage of make-whole bonds began to increase. The increase in the prevalence of make-whole bonds has been even more pronounced since the financial crisis in 2008.

	Cash	Net PP&E	Dividends
Pre-issue level	1.3184***	1.3159***	0.5549***
	(.002)	(.001)	(.006)
MW Issue	-232.565	1794.49***	-9.927**
	(734.31)	(284.05)	(3.295)
Year FE	Yes	Yes	Yes
Obs.	17,079	17,079	17,079
Within-R^2	0.9295	0.9869	0.1692

Table 31: Uses of Issuance Proceeds

Uses of Issuance Proceeds

These fixed effects regressions are constructed by regressing the accounting variable for a firm the year after it issues a bond on the variable the year before and an indicator variable for whether the bond issue was a make-whole callable issue. Thus, the coefficients for “MW Issue” should be interpreted as the change in the relevant variable after accounting for the pre-issue level relative to all other bonds. The variables are in millions of dollars, and all variables are winsorized at 1%. Robust standard errors are clustered at the industry level.

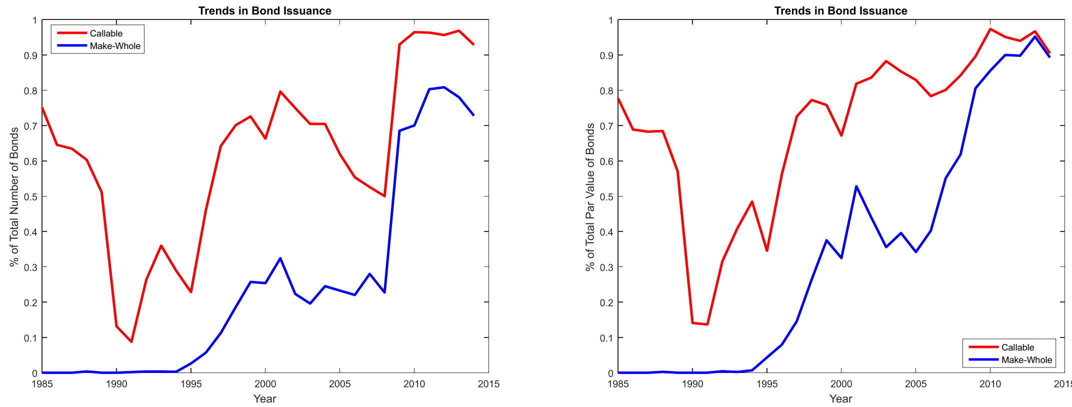

Figure 19: Percent of Total Bonds Issued by Type
These plots show the same aggregate trends as the two above. We note that callable bonds represent over 90% of both the total number of bonds and the total par value of bonds issued in the last year (and over the last five years), with the majority of this coming from make-whole bonds.

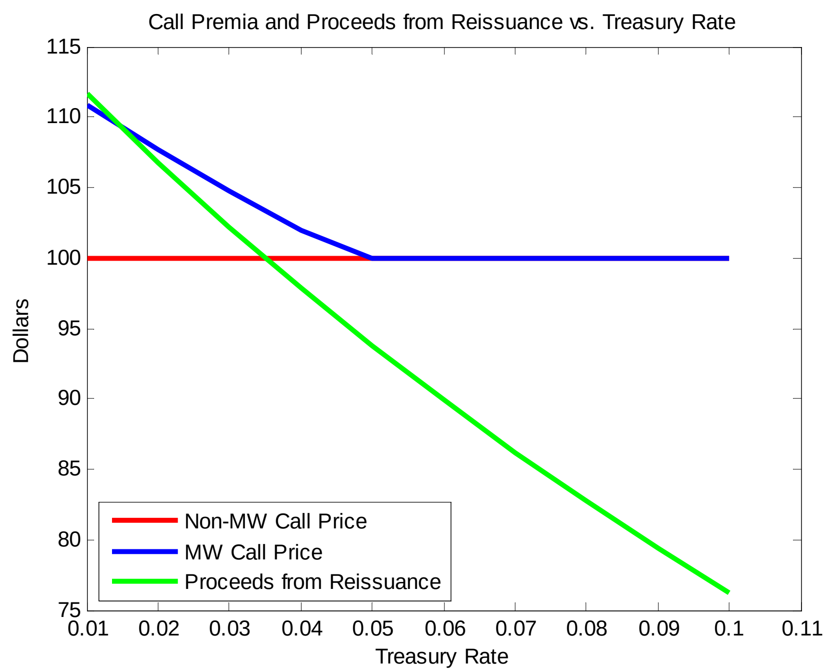

Figure 20: An Extreme Example of Make-Whole Debt
This figure replicates the example in Section 2.1 of a firm that has issued a 5-year bond at par with 5.5% annual coupon payments. The payoffs represent the proceeds/prices that the firm would have to pay after the second year of the bond, assuming a make-whole spread of 30bps. and a reissuance spread of 0bps. Note that even in the case where the firm can issue at the risk-free rate, interest rates have to be quite low for the firm to profit by calling make-whole debt and reissuing an identical bond.

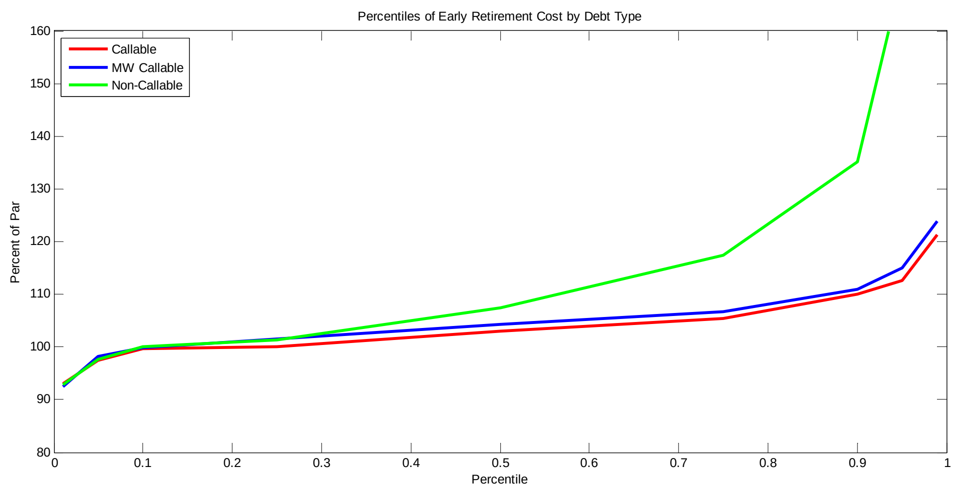

Figure 21: Costs of Early Retirement by Debt Type
The prices paid for early retirement for non-MW callable and MW callable debt are fairly similar (MW is slightly higher) and range from 90 percent of par to 110 percent of par for over 95% of the bonds retired, while the prices paid for early retirement for non-callable bonds have a greater median (by about 5% of par) and exhibit significant positive skew, suggesting that callable and make-whole bonds offer are far less risky options for early retirement.

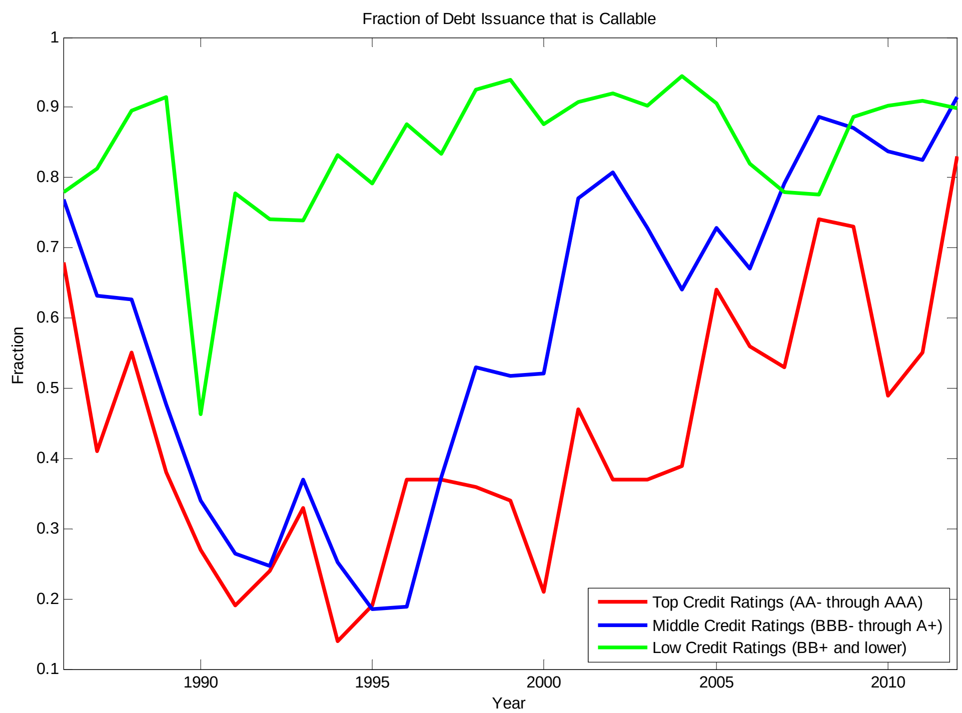

Figure 22: Callability by Firm Credit Rating
Firms with lower credit ratings (more credit risk) are more likely to issue their debt as callable and were quicker to adopt callable debt. This trend is similar for make-whole callable debt and is fairly robust across different measures of credit risk (such as leverage and income volatility) and breakpoints of credit risk categories.

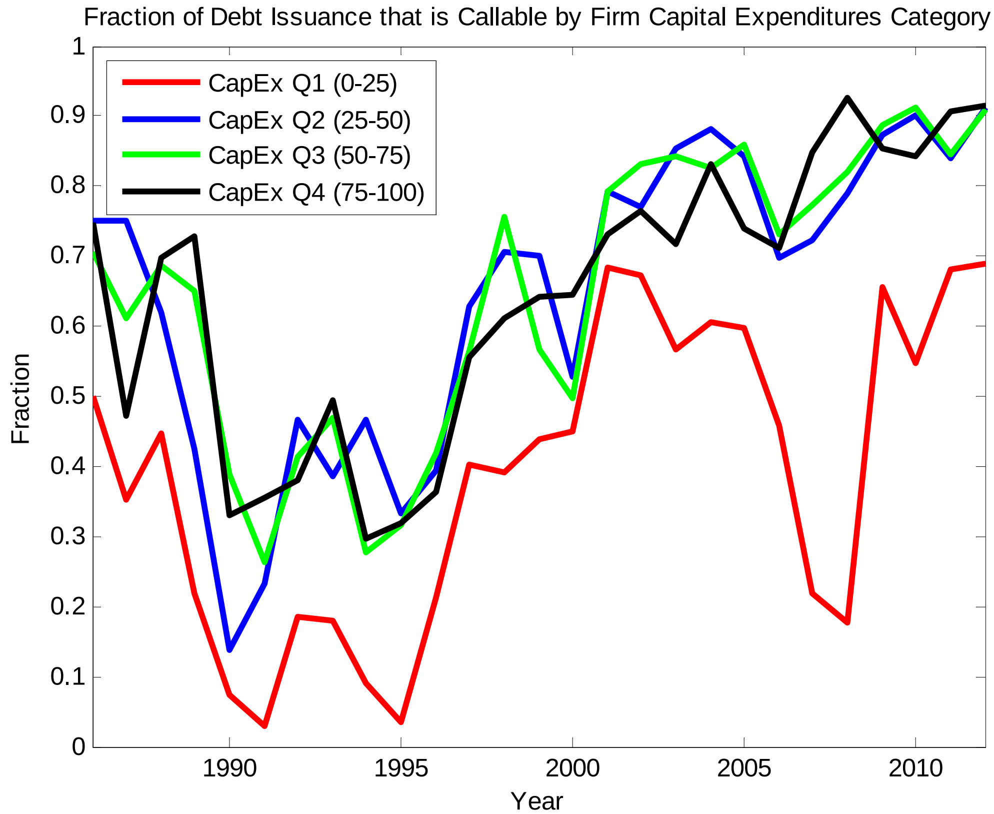

Figure 23: Callability by Firm Investment Level
Firms with higher ratios of investment to operating income are more likely to issue their debt as callable and their use of call provisions is less sensitive to market conditions. This trend is similar for make-whole callable debt and is fairly robust across different ratios for investment. Note that it also controls for changes in the average ratios of investment to operating income since the firms are re-sorted each year.

CHAPTER 3 : **Firm Volatility and the Composition of Investment**

3.1. **Introduction**

This paper focuses on the central question of when firms pursue investment in research and development (R&D) and physical capital (measured by firm capital expenditures.) Given the relative lack of consensus in the literature, much attention is paid to the former. R&D is becoming an increasingly important investment activity for private firms, with the real R&D investment by nongovernment institutions increasing tenfold over the last 50 years. While government agencies have been reducing their investments in R&D relative to total output, private firms have significantly increased their share of R&D investment, from about 30% of total R&D spending in the mid-1960s to over 70% today. Manufacturing firms alone now spend over $100 billion in R&D annually, more than the (inflation-adjusted) R&D spending of the entire public and private sectors 40 years ago. Clearly, it is important to understand why and under what conditions firms choose to invest into research and development.[1]

One such condition that is important to understand is the aggregate business cycle. Work on the question of how corporate investment should respond to business cycles goes back as early as the seminal work of Schumpeter's 1939 work on business cycles, in which he suggests that the lower opportunity cost of investment during economic downturns should motivate increased investment during this time–that is, that firms' investment policies should be countercyclical. Hall's 1991 work expands upon this by considering the problem of a firm with a fixed labor force to be allocated between current production and the creation of organizational capital. If we think of the latter as being the sort of operational and productivity improvements generated by R&D, then Hall too predicts a countercyclical R&D investment

[1]Statistics from 2005 CBO Report "R&D and Productivity Growth"

policy–in bad times, the rewards from current period production will be lower, and hence the lost revenues from shifting productive capacity to the creation of organization capital will be smaller. Hence both Schumpeter and Hall ultimately predict a negative relation between aggregate business cycles and firm investment.

These theoretical predictions have created something of a puzzle in the R&D literature, however, as subsequent empirical work suggests that aggregate R&D, in fact, comoves positively with business cycles. Work by Fatas in 2000, Walde and Woitek in 2004, Comin and Gertler in 2006, and Barlevy in 2007 all make this point. These analyses have also showed that this holds across a broad spectrum of settings, with these papers differing in their cross-section of countries studied, measures of R&D employed, and length of impact used. While there is a bit of literature suggesting that R&D is in fact countercyclical (Rafferty and Funk's 2004 paper shows that R&D expenditures decrease in response to demand shocks), the majority of the empirical evidence seems to suggest positive comovement between the R&D expenses of the private sector and aggregate business cycles.

As a response to this, several authors have recently suggested models that generate procyclical R&D patterns. Barlevy's 2007 paper generates this by relying on short-termism by entrepreneurs, and Francois and Lloyd-Ellis's 2009 paper adopt Schumpeter's framework but add long-term uncertainty to the firm's problem.

We take a different approach to this puzzle by focusing not on the comovement of aggregate R&D with aggregate business cycles, but rather by studying how firm-specific investment measures comove with firm-specific profitability. In that respect, both the focus and the results of this paper are more similar to the literature on the dispersion in investment cyclicality. For example, Ouyang's 2008 paper analyzes the comovement between R&D spending and industry-level demand shocks and finds significant cross-sectional differences in how industries change their R&D spending in response to demand shocks. She suggests that liquidity constraints and technological differences account for some of this heterogeneity. This heterogeneity is also a theme of studies that have looked at how R&D moves with aggregate business cycles, with Aghion et. al. (2012) finding that R&D is countercyclical for firms without credit constraints but becomes more procyclical as credit constraints bind more. Similarly, Lopez-Garcia et. al. (2012) find that increased internal resources make R&D less procyclical and Arvanitis et. al. (2015) find that R&D is more procyclical for firms that face higher price competition.

This paper differs from these previous studies in at least two important ways. First, rather that focusing on the comovement with aggregate or industry-specific business cycles, this paper studies the comovement between firm-specific investment and firm-specific productivity. This allows us to more clearly identify the heterogeneity across firms and to remove the confounding effect of differing exposures to aggregate conditions. Second, in addition to studying R&D and capital expenditures in isolation, this paper also studies the composition of investment (i.e. the ratio of R&D to the sum of R&D and capital expenditures) and shows that this is an economically significant quantity. Figure 1 plots this ratio over time with the aggregate business cycle, but the majority of the other analyses in the paper deal with firm-specific profitability.

Another contribution of this paper is that, in addition to demonstrating the links between firm profitability and investment composition, we also show the link between the composition of investment and the future volatility of the firm's income and equity returns. Using this, and after a discussion of why the new patterns of investment do not necessarily fit existing models, we suggest a new mechanism for the timing of investment. This mechanism draws heavily on the links between investment and future volatility and is meant simply as a guide for future models. We also display preliminary results for how the composition of investment relates to future equity returns, building upon the work of Chan, Lakonishok, and Sougiannis (2001) and Lev and Sougiannis (1996).

The remainder of the paper is organized as follows. Section 2 discusses the data used for the analyses in this paper and details some of the methodology employed to construct quantities of interest. The empirical results of interest are presented in section 3, along with a brief discussion of how these results compare to predictions of earlier theories. Section 4 presents a possible mechanism for the observed behavior of investment, drawing upon the results in section 3 and briefly discusses how this might be extended to capture evidence on equity returns. Section 6 concludes.

3.2. **Data and Methodologies**

The data from this project comes primarily from the CRSP/Compustat Merged dataset on WRDS. Compustat observations are collected on an annual basis to reduce any effects of seasonality, and merged with CRSP based on the permanent "permco" link between the two. In addition to restricting attention to the Compustat observations that could be merged with CRSP, the SIC codes 0-999 (agriculture, fishing, hunting), 4900-4999 (utilities), 6000-6999 (financials), 8888 (foreign governments), and 9000-9999 (international/non-operating) were eliminated. Companies that did not report asset, profitability, capital expenditure, or R&D levels were also eliminated, with the R&D elimination reducing the dataset by about half (from 225,134 observations that satisfied the asset, profitability and capex requirements to 117,388 that satisfied all four.) All Compustat variables were winsorized at the 1% level.

For the volatility analysis, CRSP daily returns were used, and a procedure similar to that in Kelly, Lustig, et. al. (2015) was employed. First, for all daily equity returns, a CAPM model was employed and so these returns were regressed on the returns of a daily value-weighted market index. From this we obtained the systematic portion of returns and constructed the idiosyncratic component of returns as the difference between the total and systematic returns on a daily basis for all stocks in the sample. Unlike in Kelly, Lustig, et. al., we also allowed for time-varying equity betas on an annual basis. The systematic and idiosyncratic volatilities were then constructed as the annual standard deviations of the daily systematic and idiosyncratic returns, respectively. It is also important to note here that since daily returns were used, the results presented are robust to one's choice of an equity return model. Specifically, since the Fama-French 3-factor model and principal components models do not explain significantly more variation of daily returns, the systematic and idiosyncratic components of returns are fairly similar under those constructions.

After merging these datasets, the final dataset contained 55,273 observations ranging from 1971 to 2013. This dataset also contains the annual equity returns and Fama-French factors (from Kenneth French's webpage) used for the Fama-MacBeth procedure discussed at the end of section 4. This data was used for all studies involving both equity return or volatility data and firm-specific variables. For analyses involving only firm-specific variables, the pre-merge Compustat data was used to give access to a larger panel of data. Those results are robust to use of the final merged dataset. Also note that the Compustat data differs slightly from the aggregate NSF data that some use to measure R&D, but as Figure 2 of Barlevy (2007) shows, the difference between the two is not hugely significant on an aggregate level.

3.3. **Empirical Results**

In this section we present empirical results on two significant questions. First, we study the relationship firm profitability and the composition of firm investment. Second, we show how the composition of investment translates into future equity and cash flow volatility.

3.3.1. Cyclicality

We begin by examining the relationship between a firm's profitability and what forms of investment the firm pursues. In order to answer this question, we examine three primary variables: the year-on-year revenue growth of the firm, the R&D expenses in the following year, and the capital expenditures (Capex) in the following year. In order to measure the rate of investment, we would ideally like to scale these measures by an aggregate stock of investment. As a proxy for this we scale both annual investment measures by the net stock of property, plant, and equipment (PPE) at the beginning of the period. The results that follow are robust to other scaling measures and also to simply using growth rates of Capex and R&D (although more intuitively appealing, the lumpiness of spending on these measures–particularly R&D–made scaling a more attractive solution.) Since all of this data comes from Compustat, we employ the larger Compustat-only dataset discussed in the

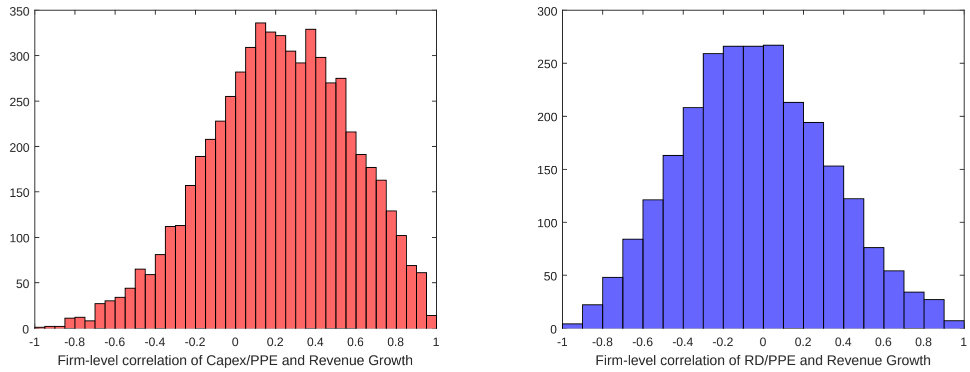

Figure 24: Firm-Level R&D and Capex Cyclicality

previous section.

The following two histograms present the frequencies of firm-level correlations between revenue growth from period t-2 to t-1 and investment spending in period t for both the Capex and R&D measures of investment. Only firms with at least 10 years of data are considered:

There are several points which are immediately apparent from these distributions. First, there exists significant cross-sectional heterogeneity in even the comovement of firm-specific profitability and future investment. The standard deviations for both measures are approximately 0.35, which is quite significant given the bounded range of the correlations. This indicates that the heterogeneity found in earlier papers is not merely due to differing exposures to aggregate or industry conditions, but rather also reflects firm-specific characteristics. We test the specific hypothesis of this being due to access to credit later in the section.

Second, we see that Capex comoves more positively with firm-specific profitability than does R&D. Specifically, 73.2% of firms exhibit a positive correlation between revenue growth and future (next year) Capex spending, while the majority–55.7%–of firms have a negative correlation between revenue growth and next year R&D expenditures. For firms in this sample, the median correlation between revenue growth and future Capex was 0.2245, while

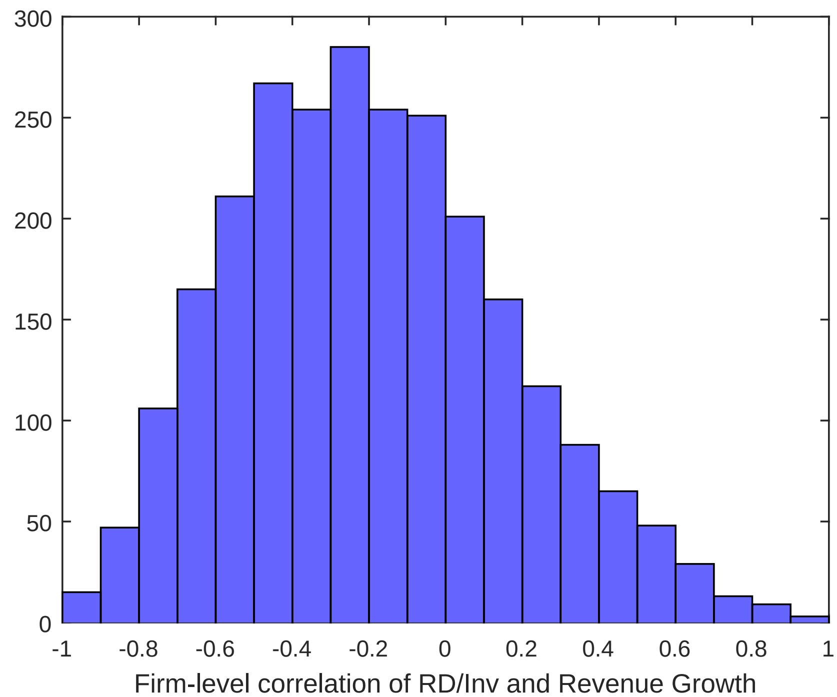

Figure 25: Firm-Level R&D/Capex Cyclicality

the median correlation for R&D was -0.05.

We can see this second point more clearly be examining the ratio of R&D expenses to total investment, which is again defined as the sum of Capex and R&D. The figure below replicates the methodology used to generate the previous histograms for this ratio:

As expected, we see that the majority of firms have negative correlations between revenue growth and the future ratio of R&D to total investment–implying that this ratio tends to decrease after positive future revenue growth. Thus, following future revenue growth, the majority of firms in this sample tend to increase Capex more than R&D in the following year. This is the case for 71.7% of firms in this sample, and we note that the average correlation coefficient between R&D/Investment and past revenue growth is -0.2. A simple regression illustrates this point quite clearly:

Note: Fixed-effects panel regression estimates. *** indicates significant at the 1% level.

	$\frac{RD_{t+1}}{Inv_{t+1}}$
ΔRev_t	-0.0392*** (-12.30)
$\frac{RD_t}{Inv_t}$	0.4940*** (132.24)
Firm Controls	Leverage, Age, Size
Firm FE	Yes
Year FE	Yes
Obs.	59,443
$Within - R^2$	31.38
$Overall - R^2$	85.03

Table 32: Regression results for R&D/Investment

Inv defined as the sum of capital expenditures and R&D expenses. Leverage defined as debt/assets, age as time since induction into Compustat database, and size as ln(assets). Standard errors are clustered at the firm level.

We see that current period revenue growth very significantly negatively predicts the future fraction of R&D to total assets. In particular, a one standard deviation increase in revenue growth today implies a 4% decrease in the ratio of R&D expenses to total investment next period, after conditioning upon that ratio this period and firm controls (as well as firm and year fixed effects.)

Now we return to another question raised by the first set of histograms: identifying what is generating the significant dispersion in the comovement of revenue growth with future investment levels. In particular, we wish to examine the hypothesis that borrowing constraints of some form are responsible for this dispersion. We use firm leverage as a proxy for borrowing constraints in this example to maintain usage of the Compustat data and keep the cross-section as large as possible. In the results (the full version of which are presented in figure 2 in the appendix), we see that leverage is a significant predictor of the comovement of investment with both firm-specific profitability conditions and aggregate business cycles. In particular, we see that higher leverage implies a higher comovement between profitability/business cycles and investment, consistent with the evidence in Aghion (2012) and others. We note, however, that leverage only explains a small part of the cross-sectional heterogeneity in these quantities. (Given how low the R-squared values are, this is likely a consequence both of leverage being an imperfect measure of a firm's access to capital and a firm's access to capital being only part of the story.)

Finally, a point that bears mentioning is that some have argued that R&D is less procyclical than Capex simply because it requires more persistent expenditures to maintain the capital. This explanation holds that without consistent R&D investment, the accumulated R&D capital will depreciate quickly, and that this is not the case for capital expenditures and physical capital. While we do find some empirical evidence for this–the stock of in-process R&D is a significant predictor for future R&D expense, differences in the stocks of R&D only explain 6% of the cross-sectional variation in future R&D expenses, indicating that there are other factors that differ across firms and which impact firms' decisions to invest in R&D.

3.3.2. Volatility

In order to explore one such possible factor, we first present evidence on how the composition of investment affects the future equity and cash flow volatility of firms. For this analysis we return to the merged CRSP-Compustat dataset and examine how investment in period t impacts the annual idiosyncratic and systematic equity volatility levels over the following year and future years. We also examine how investment in period t impacts future cash flow volatility. For both of these analyses, the investment measures used are the ratios of R&D and Capex to total assets, primarily since this measure is slightly less volatile and these regressions control for asset levels. Again, the results presented are robust to different scaling factors and to using growth rates.

The first table presents the impact of investment in period t on the total, systematic, and idiosyncratic volatility levels over the following year:

Note: Fixed-effects panel regression estimates. *** indicates significant at the 1% level.

	$TotalVol_{t+1}$	$SysVol_{t+1}$	$IdioVol_{t+1}$
$\frac{RD_t}{Assets_t}$	0.0189***	-0.0002	0.0200***
	(5.33)	(-0.12)	(5.77)
$\frac{Capex_t}{Assets_t}$	0.0094***	0.0039***	0.0091***
	(4.95)	(5.38)	(4.65)
$IdioVol_{t-1}$	0.1886***	0.0079***	0.1923***
	(5.96)	(2.65)	(6.03)
$SysVol_{t-1}$	-0.2136***	0.1176***	-0.2864***
	(-10.42)	(16.56)	(-14.18)
Firm Controls	Leverage, Age, Size	Leverage, Age, Size	Leverage, Age, Size
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	53,386	53,386	53,386
$Within-R^2$	4.48	3.76	5.88
$Overall-R^2$	27.48	13.88	33.06

Table 33: Regression results for t+1 Volatility

Leverage defined as debt/assets, age as time since induction into Compustat database, and size as ln(assets). Standard errors are clustered at the firm level.

There are several interesting features to notice from these results. First, note that RD/Assets and Capex/Assets are normalized by the same amounts so that the regression coefficients correspond to equal dollar amounts of increased investment. Beginning with the leftmost column, we see that an increase in R&D increases the future volatility of equity returns by approximately double the amount of an equivalent increase in Capex, and that both are very significant. This increase comes almost exclusively from the effect on idiosyncratic volatility–R&D does not seem to affect future systematic volatility at all. Capex does significantly impact future systematic volatility, but the primary source of increased volatility for both is through idiosyncratic volatility. (The decomposition of daily returns was such that idiosyncratic volatility is approximately 4 times higher than systematic volatility.) For idiosyncratic volatility, R&D again has about double the impact on future volatility than does Capex.

We also see that this effect is persistent. The following table reproduces the first plot for equity volatility two years from the investment:

	$TotalVol_{t+2}$	$SysVol_{t+2}$	$IdioVol_{t+2}$
$\frac{RD_t}{Assets_t}$	0.0207***	0.0012	0.0213***
	(4.96)	(0.62)	(5.26)
$\frac{Capex_t}{Assets_t}$	0.0091***	0.0091***	0.0086***
	(4.33)	(3.25)	(4.09)
$IdioVol_{t-1}$	0.0057**	-0.0163***	0.0674**
	(2.14)	(-5.08)	(2.44)
$SysVol_{t-1}$	-0.1792***	0.0687***	-0.2344***
	(-8.42)	(9.16)	(-10.89)
Firm Controls	Leverage, Age, Size	Leverage, Age, Size	Leverage, Age, Size
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	46,776	46,776	46,776
$Within - R^2$	1.67	3.22	2.63
$Overall - R^2$	13.11	2.35	21.21

Table 34: Regression results for t+2 Volatility

Note: Fixed-effects panel regression estimates. *** indicates significant at the 1% level, ** indicates significant at the 5% level. Leverage defined as debt/assets, age as time since induction into Compustat database, and size as ln(assets). Standard errors are clustered at the firm level.

The results here are very similar to those in the previous table. Namely, a dollar invested in R&D increases future total and idiosyncratic volatility more than two times as much as an equivalent dollar invested in Capex. R&D investments do not seem to have a significant effect on future systematic volatility, but Capex does. Moreover, this effect is stronger two years out, so that Capex increases systematic volatility by slightly more than idiosyncratic volatility after two years.

While equity volatility offers the benefit of being easy to measure, it also is potentially affected by many factors, not least market risk exposures and firm-specific leverage. In order to help determine whether these factors are biasing the results we present, we also run the same panel regressions using the volatility of asset levels and return on assets (ROA), defined as net income divided by assets. These volatility measures are computed by taking the standard deviation of these measures over the four quarters following the investment. We also note that the net income measure is robust to other definitions of profitability and other denominators. In the full results, presented in figure 3 in the appendix, we see that, although R&D and Capex both lead to similar increases in future asset levels, a dollar invested in R&D increases future ROA volatility by almost triple the effect of a dollar increase in Capex. These results confirm our equity volatility results and are also consistent with previous estimates in the accounting literature (see e.g. Kothari et. al. (2002)).

Thus, it seems fairly clear that R&D has a larger effect on future equity and cash flow volatility than does an equivalent investment in Capex. This seems to come mainly from the effect on idiosyncratic volatility, since R&D has no significant effect on future systematic volatility.

3.4. **Potential Mechanism**

From these empirical results we see that R&D seems to comove more negatively with firm-specific profitability than Capex and that R&D tends to be tied to greater future firm volatility–particularly equity volatility. We also see that, while explanations relating to credit constraints and the necessity of maintaining R&D capital may account for some of the motivation behind these investment decisions, they fail to have large explanatory power. Furthermore, while earlier opportunity cost theories may potentially explain the more negative comovement of R&D with profitability than Capex, we thought that it might be interesting to show how another explanation, drawing upon elements of a few of these previous theories, can link the empirical results on volatility and investment timing.

This is intended to be a simple exercise in how these facts can be linked and a basis for future theoretical work more than as a formal model in its own right, but perhaps the simple model and the intuition embedded within it can prove useful.

3.4.1. Model Outline

We consider the simplest setting possible to show the intuition. Consider a three period model in which firms can choose investment into both R&D and Capex. Firms are subject to aggregate and idiosyncratic shocks, and firms begin with an initial level of equity. We do not allow firms to issue equity beyond their initial levels, so any investment above their initial equity level must be financed by debt, which is endogenously priced by risk-neutral debtholders at an interest rate r (we assume some nonzero cost of financial distress.) Firms can differ at time 0 in at most two dimensions: their initial idiosyncratic state and their level of initial equity. The idiosyncratic productivity X_t and the aggregate productivity Z_t follow the following processes:

$$log(X_t) = x_t = (1 - \rho_x)\mu_x + \rho_x x_{t-1} + \sigma_x \epsilon_{xt}$$

$$log(Z_t) = z_t = (1 - \rho_z)\mu_z + \rho_z z_{t-1} + \sigma_z \epsilon_{zt}$$

where ϵ_{xt} and ϵ_{zt} are drawn from independent standard normal distributions.

At time 0, the firm observes the initial values of both productivity processes and then chooses levels of investment in R&D R_0 and Capex C_0. This investment is financed first with the initial equity that the firm has, and then with debt. Next period, the firm makes no decision, but the stock of capital that the firm has becomes $K_1 = X_1 R_1 + Z_1 C_1$. Thus here we envision R&D as being research that potentially (dependent on the firm's productivity) can turn into productive capital, but possibly may not. We also simplify the empirical results a bit and have idiosyncratic volatility depend only on the amount of R&D invested and systematic volatility depend only on the amount of Capex invested. In a more sophisticated model one could allow both types of volatility to weight differentially on both types of investment.

Lastly, in the third period a firm earns $Z_2 K_1^{\alpha}$, where $\alpha < 1$ represents decreasing returns to scale. The firm's debt is due, and it can sell its full stock of capital to pay back its debt. We assume no depreciation, which is again something that can be relaxed to perhaps more closely capture the idea that R&D capital depreciates faster than physical capital. The firm's resulting equity value can be given by the sum of its profit, capital, and uninvested equity less debt (and taxes π.)

Formally, the firm solves the following problem:

Firm chooses $I_{R,0}, I_{C,0}$ such that:

$$(R_0, C_0) = \max_{(IR_0', C_0')} E_0 \left[\beta max \left\{ \left[ZK_1^{\alpha} \left(R_0', C_0' \right) - r \left(R_0', C_0' \right) D_0 \right] (1-\tau) - NetDebt + K_1 \, , \, 0 \right\} \right]$$

where:

$$K_1 = X_1 R_0' + Z_1 C_0'$$

$$D_0 = max \left\{ 0 \, , \, R_0' + C_0' - E_0 \right\}$$

3.4.2. Model Intuition and Comparative Statics

The key to this mechanism is that debt is endogenously priced in the model. Since returns from Capex vary more with aggregate business cycles than returns from R&D, the persistence of aggregate business cycles will imply the the expected change in equity value from a marginal Capex investment is greater when the current aggregate business cycle is in a good state. This effect will be greater for Capex than R&D simply because future Capex returns are more exposed to aggregate business cycles. These higher expected future cash flows imply lower probability of default and thus lower costs of debt financing for firms.

What constrains this is decreasing returns to scale. Firms partially trade off Capex and R&D investment due to both the decreasing returns to scale and the fact that they need to finance investment with debt. While R&D also has a lower cost of financing during positive business cycles, it benefits less since its cash flows are less directly tied to aggregate conditions, implying that it will be less procyclical than Capex. Furthermore, since the returns to R&D depend in large part on an uncorrelated idiosyncratic productivity, firms will be incentivized to use internal capital and possibly issue debt to finance R&D in cases

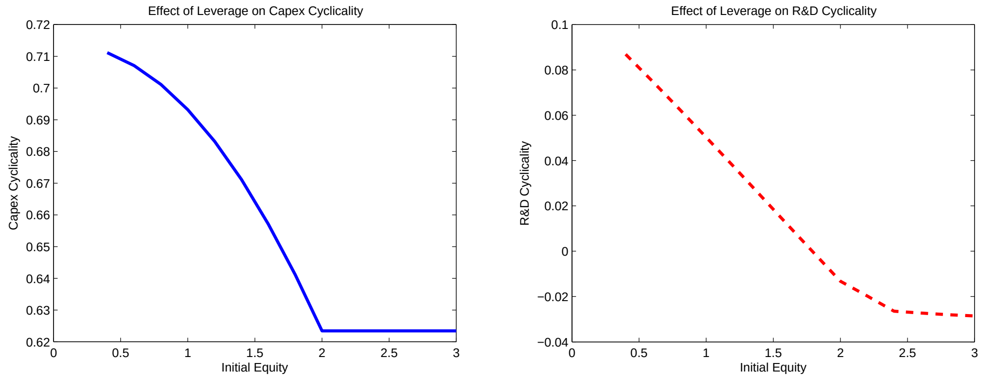

Figure 26: Model-Implied Effect of Leverage on Investment Cyclicality

in which they have high idiosyncratic productivity, even if the aggregate state is low. Thus, Capex will be significantly more procyclical than R&D with respect to aggregate conditions. Since firm profitability is driven primarily by aggregate conditions, this effect will also drive Capex to be more positively associated with firm profitability than R&D. A preliminary numerical exercise with this model showing some of these effects is given in figure 4 in the appendix.

Another feature of the model which matches empirical evidence is that firms with lower leverage will have investment policies which covary less with aggregate business cycles and with firm-specific profitability. The following plots show how increasing the initial equity (and hence reducing the leverage) reduces the covariance of both Capex and R&D with aggregate conditions (under the baseline parameters in figure 4.) Similar results also obtain for the covariance with firm profitability.

3.5. **Conclusion**

We began this paper by demonstrating several novel empirical trends that attempt to address the puzzle of why aggregate R&D seems procyclical, even when the most highly recognized models of investment timing suggest that it should be countercyclical. In particular, we showed that when one looks at the comovement of firm profitability (measured by revenue growth) with future firm R&D expenditures, the relation is weakly negative.

Furthermore, when compared to the comovement of firm profitability with future Capex, R&D certainly responds far more negatively. The cross-sectional patterns also lent some credence to existing theories about the irreversibility of R&D flows and the effect of credit constraints, but it seemed that there was more to the story.

To explore this, we also documented another new empirical fact: that R&D impact future equity volatility almost twice as much as Capex, and that this comes almost solely through the effect on idiosyncratic volatility. While Capex increases both systematic and idiosyncratic equity return volatility, R&D has no significant effect on the systematic component on volatility. We also showed that this effect was even stronger for cash flow volatility.

Next, we demonstrated how these volatility effects could cause the patterns in investment composition observed in the data. In order to achieve this result, the trade-off between forms of investment enforced by the decreasing returns to scale of the production function and the endogenous pricing of debt and limited internal capital were important, as were the different loadings on aggregate and idiosyncratic productivity of the two types of investment. We briefly another fact that this model might be extended to capture: that Fama-French 3-factor alphas increase monotonically with the portion of investment devoted to R&D.

Ultimately, it is hoped that this paper will have provided several new facts about the composition of investment, the relation between the composition of investment and equity and cash flow volatility, and the link between equity alphas and investment composition. We demonstrated one possible link between these facts, but there are surely others to be explored in the future.

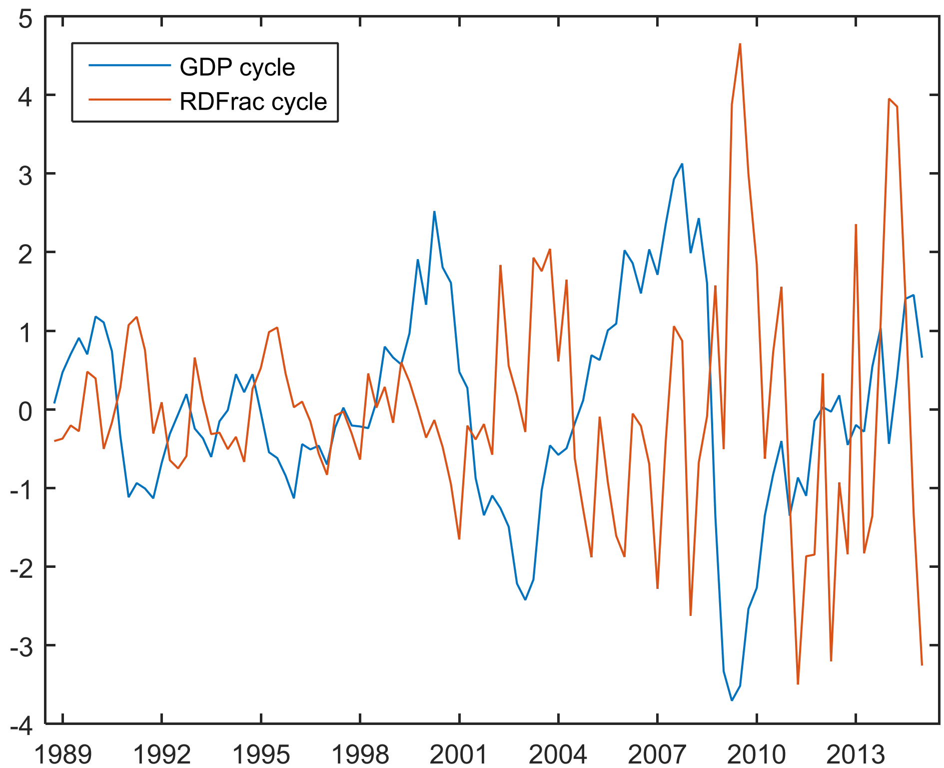

Figure 27: Aggregate R&D Fraction
This figure plots the cyclic component of the aggregate fraction of R&D to total investment with the cyclic component of GDP. We see the the two measures have a negative relation, with a correlation coefficient of -0.36, implying that Capex is more procyclical than R&D on aggregate. This negative relation is robust to using growth rates of these quantities as well.

3.6. **Appendices and Figures**

Aggregate R&D Fraction, 1989-2014

	$\rho(\frac{Capex}{EBITDA}, Revcycle)$	$\rho(\frac{RD}{EBITDA}, Revcycle)$	$\rho(\frac{Capex}{PPE}, GDPcycle)$	$\rho(\frac{RD}{PPE}, GDPcycle)$
Leverage	0.3841***	0.2492**	0.2399**	0.2672**
	(5.03)	(2.09)	(3.22)	(2.16)
Firm Controls	Yes	Yes	Yes	Yes
R^2	4.04	3.08	3.16	1.65

Table 35: Effect of Leverage on Investment Cyclicality

Effect of Leverage

Note: Fixed-effects panel regression estimates. *** indicates significant at the 1% level, ** indicates significant at the 5% level. Leverage defined as debt/assets, age as time since induction into Compustat database, and size as ln(assets). Standard errors are clustered at the firm level. Firm controls include average levels of the following: $\frac{RD}{Assets}$, $\frac{Capex}{Assets}$, $\frac{RD}{EBITDA}$, $\frac{Capex}{EBITDA}$, $\frac{PPE}{Assets}$, $\frac{EBITDA}{Assets}$, $\frac{NI}{Assets}$, $\frac{RD}{Capex}$, $\frac{Book}{Market}$, $\frac{LTDebt}{TotalDebt}$, $Assets$. Results respond to different growth and investment measures, cycles vs. growth rates, and various control groups.

	$AssetVol_{t+1,..,4}$	$ROAVol_{t+1,..,4}$
$\frac{RD_t}{Assets_t}$	159.28***	0.0426***
	(3.58)	(6.16)
$\frac{Capex_t}{Assets_t}$	152.77	0.0147***
	(0.86)	(2.11)
$AssetVol_{t-1,..,4}$	0.0725	
	(1.20)	
$ROAVol_{t-1,..,4}$		-0.1012***
		(-8.09)
Firm Controls	Leverage, Age, Size	Leverage, Age, Size
Firm FE	Yes	Yes
Year FE	Yes	Yes
Obs.	63,199	60,857
$Within - R^2$	1.78	2.84
$Overall - R^2$	21.55	8.78

Table 36: Asset and ROA Volatility

Asset and ROA volatility

Note: Fixed-effects panel regression estimates. *** indicates significant at the 1% level. Leverage defined as debt/assets, age as time since induction into Compustat database, and size as ln(assets). Standard errors are clustered at the firm level. Asset volatility responds roughly similarly to increases in Capex and R&D, while ROA volatility increases about three times more for a dollar of R&D investment vs. a dollar of Capex investment.

Figure 4: Sketch of Model Results

- Correlations between aggregate investment levels and aggregate cycles are as follows:

	Model	Data
$\rho\left(R_0, Z_0\right)$	-0.01327	-0.02313
$\rho\left(C_0, Z_0\right)$	0.623442	0.597821
$\rho\left(\frac{R_0)}{C_0+R_0}, Z_0\right)$	-0.34229	-0.36

- While some claim a positive correlation between aggregate R&D and GDP cycles, we found a slightly negative one. The model, however, can match a positive correlation between aggregate R&D and business cycles. (For example, by increasing the implied credit constraints by reducing the implied leverage to 0.5, the implied correlation between R&D and aggregate states increases to approximately 0.15, while the correlation between aggregate capex and the aggregate cycle increases marginally to 0.64.

- Model parameters are as follows:

- Standard Values

  - $\beta = 0.98$ –matching average annual risk-free rate

  - $\tau = 0.4$ –U.S. corporate tax rate

  - $rr = 0.6$ –recovery rate in default

- Productivity Processes

  - Idiosyncratic: $\overline{X}_t = 1.03$, $\sigma(X_t) = 0.06$, $\rho(X_t) = 0.3$

  - Aggregate: $\overline{Z}_t = 1.04$, $\sigma(Z_t) = 0.025$, $\rho(Z_t) = 0.6$

- Parameters used to target data moments

- $E_0 = 2$ –initial equity
    * relative to $R_0 \in [0, 2]$, $C_0 \in [0, 2]$
- $\alpha = 0.6$ –decreasing returns to scale parameter

BIBLIOGRAPHY

P. Aghion, P. Askenazy, N. Berman, G. Cette, and L. Eymard. Credit constraints and the cyclicality of r and d investment: Evidence from france. *Journal of the European Economic Association*, 10(5):1001–1024, 2012.

H. Ai, M. Croce, and K. Li. Towards a quantitative equilibrium model of intangible capital. *Review of Financial Studies*, 26:491–538, 2013.

Y. Ait-Sahalia, J. A. Parker, and M. Yogo. Luxury goods and the equity premium. *The Journal of Finance*, 59:2959–3004, 2004.

H. Almeida, M. Campello, B. Laranjeira, and S. Weisbenner. Corporate debt maturity and the real effects of the 2007 credit crisis. Technical report, National Bureau of Economic Research, 2009.

S. Arvanitis, B. Lokshin, P. Mohnen, and M. Worter. Impact of external knowledge acquisition strategies on innovation: A comparative study based on dutch and swiss panel data. *Review of Industrial Organization*, 2015. forthcoming.

C. S. Asness, A. Frazzini, and L. H. Pedersen. Quality minus junk. Working paper, AQR Capital, 2014. URL `https://ssrn.com/abstract=2312432`.

J. C. Banko and L. Zhou. Callable bonds revisited. *Financial Management*, 39(2):613–641, August 2010.

G. Barlevy. On the cyclicality of research and development. *American Economic Review*, 97(4):1131–1164, 2007.

A. Barnea, R. A. Haugen, and L. W. Senbet. A rationale for debt maturity structure and call provisions in the agency theoretic framework. *The Journal of Finance*, 35(5): 1223–234, November 1980.

Z. Bodie and R. Taggart. Future investment opportunities and the value of the call provision on a bond. *Journal of Finance*, pages 1187–1200, September 1978.

M. K. Brunnermeier and M. Yogo. A note on liquidity risk management. Technical report, National Bureau of Economic Research, 2009.

B. Cassiman and S. Vanormelingen. *Profiting from Innovation: Firm Level Evidence on Markups*. Centre for Economic Policy Research, 2013.

A. Chakraborty and B. Yilmaz. Adverse selection and convertible bonds. *Review of Economic Studies*, 78:148–175, May 2011.

L. K. C. Chan, J. Lakonishok, and T. Sougiannis. The stock market valuation of research and development expenditures. *Journal of Finance*, 56:2431–2456, 2001.

H. Chen, Y. Xu, and J. Yang. Systematic risk, debt maturity, and the term structure of credit spreads. August 2013. URL https://papers.ssrn.com/sol3/papers.cfm/abstract_id=2024329.

J. Choi, D. Hackbarth, and J. Zechner. Granularity of corporate debt. Technical report, CFS Working Paper, 2013.

L. Cohen, K. Diether, and C. Malloy. Misvaluing innovation. *Review of Financial Studies*, 26(3):635–666, 2013.

D. Comin and M. Gertler. Medium-term business cycles. *American Economic Review*, 96 (3):523–551, 2006.

A. Corhay, H. Kung, and L. Schmid. *Competition, Markups and Predictable Returns*. 2015. URL https://ssrn.com/abstract=2550981.

L. E. Crabbe and J. Helwege. Alternative tests of agency theories of callable corporate bonds. *Financial Management*, 23:4, 1994.

D. W. Diamond. Debt maturity structure and liquidity risk. *The Quarterly Journal of Economics*, 106(3):709–737, 1991.

A. Eberhart, W. Maxwell, and A. Siddique. An examination of the long-term abnormal stock returns and operating performance following r and d increases. *Journal of Finance*, 59:623–650, 2004.

A. L. Eisfeldt and D. Papanikolaou. Organization capital and the cross-section of expected returns. *The Journal of Finance*, 68:1365–1406, 2013.

E. F. Fama and K. R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

E. F. Fama and K. R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015.

A. Fatas. Do business cycles cast long shadows? short-run persistence and economic growth. *Journal of Economic Growth*, 5(2):147–162, 2000.

M. Faulkender. Hedging or market timing? selecting the interest rate exposure of corporate debt. *The Journal of Finance*, 60(2):931–962, February 2005.

P. Francois and H. Lloyd-Ellis. *Schumpeterian Cycles with Pro-cyclical R and D.* 2008. URL http://www.mcgill.ca/les/.

K. A. Froot, D. S. Scharfstein, and J. C. Stein. Risk management: Coordinating corporate investment and financing policies. *The Journal of Finance*, 48(5):1629–1658, 1993.

B. Glover. *The Expected Cost of Default.* The Wharton School, University of Pennsylvania, 2013. Thesis.

J. F. Gomes and L. Schmid. Equilibrium credit spreads and the macroeconomy. September 2010. URL http://ssrn.com/abstract=1363840.

J. R. Graham and C. R. Harvey. The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics*, 60:187–243, 2001.

R. Greenwood and S. G. Hanson. Issuer quality and corporate bond returns. *Review of Financial Studies*, 26(6):1483–1525, 2013.

G. M. Grossman and E. Helpman. *Innovation and Growth in the Global Economy.* MIT Press, Cambridge, Mass., 1997.

L. Gu. Product market competition, r and d investment, and stock returns. *Journal of Financial Economics*, 119:441–455, 2015.

L. Guntay, N. R. Prabhala, and H. Unal. Callable bonds and hedging. 2002. URL https://ideas.repec.org/p/wop/pennin/02-13.html.

B. H. Hall and J. Lerner. Chapter 14 - the financing of r and d and innovation. In *Handbook of the Economics of Innovation*, pages 609–639. North-Holland, 2010.

R. Hall. Recessions as reorganizations. *NBER Macroeconomic Annual*, 1991.

Z. He and G. Matvos. Debt and creative destruction: Why could subsidizing corporate debt be optimal? March 2012. URL http://www.nber.org/papers/w17920.

Z. He and K. Milbradt. Endogenous liquidity and defaultable bonds. Technical report, National Bureau of Economic Research, 2012.

Z. He and W. Xiong. Rollover risk and credit risk. *The Journal of Finance*, 67(2):391–430, 2012.

B. Herskovic, B. T. Kelly, H. N. Lustig, and S. Van Nieuwerburgh. The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics*, 2014. Forthcoming.

D. Hirshleifer, P.-H. Hsu, and D. Li. Innovative efficiency and stock returns. *Journal of Financial Economics*, 107:632–654, 2013.

K. Hou, C. Xue, and L. Zhang. A comparison of new factor models. Working paper, Ohio State University, Charles A. Dice Center for Research in Financial Economics, 2015.

U. J. Jermann and V. Z. Yue. Interest rate swaps and corporate default. February 2014. URL https://ideas.repec.org/p/ecb/ecbwps/20131590.html.

B. Julio, W. Kim, and M. S. Weisbach. What determines the structure of corporate debt issues? July 2008. URL `http://www.nber.org/papers/w13706`.

B. A. Kalyman. Bond refunding with stochastic interest rates. *Management Science*, pages 171–183, 1971.

T.-H. D. King and D. C. Mauer. Corporate call policy for nonconvertible bonds. *The Journal of Business*, 73(3):403–44, 2000.

R. Kish and M. Livingston. Determinants of the call option on corporate bonds. *Journal of Banking and Finance*, pages 687–703, August 1992.

L. Kogan, S. A. Papanikolaou, Dimitris, and N. Stoffman. *Technological Innovation, Resource Allocation and Growth*. 2016. doi: http://dx.doi.org/10.2139/ssrn.2193068. URL `https://ssrn.com/abstract=2193068`.

T. P. Korkeamaki and W. T. Moore. Convertible bond design and capital investment: The role of call provisions. *The Journal of Finance*, 59(1):391–405, 2004.

S. P. Kothari, T. E. Laguerre, and A. J. Leone. Capitalization versus expensing. *Review of Accounting Studies*, 7(4):355–382, 2002.

A. Kraus. The bond refunding decision in an efficient market. *Journal of Financial and Quantitative Analysis*, 7:93–806, December 1973.

T. Kruse, T. Nohel, and S. K. Todd. The decision to repurchase debt. *The Journal of Applied Corporate Finance*, 26(2):85–93, October 2013.

H. Kung and L. Schmid. Innovation, growth, and asset prices. *The Journal of Finance*, 70: 1001–1037, 2015.

H. E. Leland. Corporate debt value, bond covenants, and optimal capital structure. *The Journal of Finance*, 49(4):1213–1252, 1994.

H. E. Leland and K. B. Toft. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance*, 51(3):987–1019, 1996.

M. Lettau and J. A. Wachter. The term structures of equity and interest rates. *Journal of Financial Economics*, 101(1):90–113, 2011.

B. Lev and T. Sougiannis. The capitalization, amortization, and value- relevance of r and d. *Journal of Accounting and Economics*, 21:107–138, 1996.

H. Levy and R. Shalev. The decision between tender offers and open market repurchases. *Web*, 2013. URL `https://en-coller.tau.ac.il/sites/nihul_en.tau.ac.il/files/media_server/Recanati/management/seminars/account/hagit.pdf`.

D. Li. Financial constraints, r and d investment, and stock returns. *Review of Financial Studies*, 24(9):2974–3007, 2011.

X. Lin. Endogenous technological progress and the cross-section of stock returns. *Journal of Financial Economics*, 103(2):411–427, 2012.

F. A. Longstaff. Are negative option prices possible? the callable us treasury-bond puzzle. *The Journal of Business*, 65(4):571–92, 1992.

P. Lopez-Garcia, J. M. Montero, and E. Moral-Benito. Business cycles and investment in intangibles: Evidence from spanish firms. Working Paper 1219, Banco de Espana, May 2012.

S. Mann and E. Powers. Determinants of bond tender offer premiums and tendering rates. *Journal of Banking and Finance*, 31(3):547–566, March 2007.

A. Mian and J. Santos. Liquidity risk, maturity management and the business cycle. March 2011. URL `https://papers.ssrn.com/sol3/papers.cfm/abstract_id=2023516`.

K. Mitchell. The call, sinking fund, and term-to-maturity features of corporate bonds: An empirical investigation. *The Journal of Financial and Quantitative Analysis*, 26(2):201, 1991.

F. Modigliani and R. Sutch. Innovations in interest rate policy. *American Economic Review*, pages 178–197, May 1966.

S. Myers. Determinants of corporate borrowing. *Journal of Financial Economics*, pages 147–175, November 1977.

S. Myers and N. Majluf. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, pages 187–221, June 1984.

C. J. Nekarda and V. A. Ramey. The cyclical behavior of the price-cost markup. Working paper, NBER, 2013. URL `https://ssrn.com/abstract=2276363`.

R. NovyMarx. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28, 2013.

C. B. Office. *R and D and Productivity Growth*. Web, 2005. URL `https://www.cbo.gov/sites/default/_les/cbo_les/ftpdocs/64xx/doc6482/06-17-r-d.pdf`.

M. Ouyang. On the cyclicality of research and development. *Review of Economic Studies*, 93(2):542–53, 2011.

M. Ouynag. *Pro-cyclical Aggregate R and D: A Comovement Phenomennon*. 2011. URL `http://www.socsci.uci.edu/mouyang/`.

R. H. Peters and L. A. Taylor. Intangible capital and the investment-q relation. *Journal of Financial Economics*, 2016. Forthcoming.

M. Rafferty and M. Funk. Applied economics. *Demand Shocks and Firm-Financed R and D Expenditures*, 36(14):1529–1536, 2004.

V. Ramey and M. D. Shapiro. Displaced capital: A study of aerospace plant closings. *Journal of Political Economy*, pages 958–992, October 2001.

S. Sarkar. Probability of call and likelihood of the call feature in a corporate bond. *Journal of Banking Finance*, 25(3):505–533, June 2001.

J. Schumpeter. *Business Cycles: A theoretical, historical and statistical analysis of the Capitalist process*. McGraw-Hill, 1939.

E. S. Schwartz and I. Venezia. The determinants of bond call premia: A signalling approach. *Journal of Financial Services Research*, 8(4):243–256, September 1994.

Q. Sun. Cash holdings and risky access to future credit. 2014. URL `https://dornsifecms.usc.edu/assets/sites/394/docs/JobMarketPaper_QiSun.pdf`.

J. S. Thatcher. The choice of call provision terms: Evidence of the existence of agency costs of debt. *The Journal of Finance*, 40(2):549–561, 1985.

J. Tirole. *The Theory of Industrial Organization*. MIT Press, Cambridge, Mass., 1988.

J. H. vanBinsbergen and C. C. Opp. *Real Anomalies*. 2016. doi: http://dx.doi.org/10.2139/ssrn.2747739. URL `https://ssrn.com/abstract=2747739`.

J. D. Vu. An empirical investigation of calls of non-convertible bonds. *Journal of Financial Economics*, 16(2):235–265, 1986.

K. Walde and U. Woitek. R and d expenditure in g7 countries and the implications for endogenous fluctuations and growth. *Economics Letters*, pages 91–97, 82 2002.

H. M. Weingartner. Optimal timing of bond refunding. *Management Science*, pages 511–524, 1967.

Q. Xu. *Early Refinancing and Maturity Management in the Corporate Bond Market*. Chicago Booth, 2016. Thesis.

L. Zhang. The value premium. *The Journal of Finance*, 60:67–103, 2005.